برنامج
دعم تأليف كتاب

جامعة الملك سعود
King Saud University

عمادة البحث العلمي
Deanship of Scientific Research

# الإحصاء والاحتمالات
## (نظرية وتطبيقات)

تأليف

أ.د. حميد عويِّد مشرف العلكه    د. عبيد جفين جليغم القحطاني

جامعة الملك سعود

دار جامعة الملك سعود للنشر
KING SAUD UNIVERSITY PRESS

١٤٤٠هـ/٢٠١٩م

بسم الله الرحمن الرحيم

# الإحصاء والاحتمالات

## (نظرية وتطبيقات)


تأليف

الأستاذ الدكتور حميد عويِّد مشرف العكله

جامعة الملك سعود

الدكتور عبيد جفين جليغم القحطاني

جامعة الملك سعود



دار جامعة الملك سعود للنشر

KING SAUD UNIVERSITY PRESS

ص.ب ٦٨٩٥٣ - الرياض ١١٥٣٧ المملكة العربية السعودية

*فهرسة مكتبة الملك فهد الوطنية أثناء النشر*
العلكه ، حميد عويد مشرف
الإحصاء والاحتمالات ( نظرية وتطبيقات ) / حميد عويد مشرف العلكه ؛ عبيد جفين جليغم القحطاني ، الرياض ، ١٤٤١هـ
٧٣٦ص؛ ٢١×٢٨سم
ردمك: ٢-٨١٢-٥٠٧-٦٠٣-٩٧٨
١- الإحصاء الرياضي ٢- الاحتمالات ( رياضيات ) أ. القحطاني ، عبيد جفين جليغم (مؤلف مشارك). ب. العنوان
ديوي ٥١٢ ١٤٤١/٢٣٩٢

رقم الإيداع: ١٤٤١/٢٣٩٢
ردمك : ٢-٨١٢-٥٠٧-٦٠٣-٩٧٨

يتقدم المؤلف بالشكر لعمادة البحث العلمي لدعمها هذا الكتاب من خلال برنامج " دعم تأليف كتاب"

صدر هذا الكتاب عن عمادة البحث العلمي ، وقد وافقت العمادة على نشره بعد استيفائه الشروط العلمية للنشر بالجامعة بخطابها رقم ٤/٦٧/١٠١٥١٥ وتاريخ ١٤٣٨/٣/٩هـ



دار جامعة الملك سعود للنشر
KING SAUD UNIVERSITY PRESS

# مقدمة المؤلفين

الحمد لله ربِّ العالمين، والصلاة والسلام على نبينا محمّد وعلى آله وأصحابه ومن اتبعهم بإحسان إلى يوم الدين. ما من شك في أنَّ الأمم تتقدّم بالعلم والمعرفة والعمل الجاد والاجتهاد، وهذا ما أوصانا الله تعالى به، فقد أمرنا سبحانه وتعالى بطلب العلم واستخدامه في اكتشاف ما خلق والبحث والتمحيص في مكوِّناتها، حيث يقول ربّنا عزَّ وجل في الآية (٣٣) من سورة الرحمن: **يا معشر الجنِّ والإنس إن استطعتم أن تنفذوا من أقطار السماوات والأرض فانفذوا لا تنفذون إلاَّ بسلطان***، ولكيلا يتمادى البشر بصنيعهم ومعارفهم ويظنون أنَّهم قادرون عليها، فقد ذكرنا الله تعالى بقوله في الآية (٨٥) من سورة الإسراء:**وما أوتيتم من العلم إلاَّ قليلا***.

كما هو واضح من عنوان الكتاب فإنَّ العلم الذي سنتناوله هو من العلوم القديمة نسبياً وعلى وجه الخصوص علم الإحصاء. في الحقيقية إنَّ تاريخ **علم الإحصاء** غير مُحدَّد تماماً، فليس هناك ما يشير إلى زمن ولادته بشكل دقيق، فالبعض يَذكر أنَّه كان حوالي عام 1662، والبعض الآخر يذكر أنَّه يعود إلى حوالي عام 1749 . لكن في الواقع يُعتقد أنَّ علم الإحصاء قديم قِدَم تاريخ الحساب عندما بدأت المراحل الأولى للتّجمعات البشريّة، وبعد الانتهاء من مرحلة المشاع، حيث سادّت فيها مظاهر السلطة وحبّ التملُّك التي استوجبت العدَّ والتصنيف للممتلكات، والأنعام، والجند، ولذلك يُنظر إلى الإحصاء على أنَّه توأم الحساب أو ربما كان أحدهما هو الشقيق الأكبر للآخر.

في الواقع إنَّ استخدام تعبير **"الإحصاء"** بين النّاس كان متداولاً منذ زمن بعيد جداً وبدلالات واضحة، فعلى سبيل المثال كانت كلمة الإحصاء دارجة في اللغة العربية منذ أكثر من 1500 عام على الأقل ولاّ أدلَّ على ذلك من ورود كلمة إحصاء في القرآن الكريم في مواضع عديدة (**حيث كانت اللغة العربية في أوج ازدهارها**)، فقد جاءت في القرآن الكريم بمعنى سلوك العدّ حيث يقول الله تعالى في الآية (١٨) من سورة النحل: **وإن تعدّوا نعمة الله لا تحصوها***، وفي موضع آخر جاء الإحصاء بمعنى مُميّز عن مفهوم العدّ حيث يقول الله تعالى في الآية (٩٤) من سورة مريم: **أحصاهم وعدّهم عدّاً***، فلو كان لكلمة الإحصاء والعدّ المدلول نفسه لما أتيتا بهذا الترتيب والتأكيد على كلمة العدّ ب عدّاً. إذن فما هو المقصود بكلمة أحصاهم (**ومنها يمكننا فهم معنى الإحصاء**) هنا إن لم يكن العدّ؟

إنّ المقصود هنا (**والله أعلم**) هو معرفة سلوك وخصائص من وقع عليهم العدّ، وهذا ما سنلمسه بوضوح لدى دراسة الإحصاء كعلم رياضيّاتي إذ إنّنا سنلاحظ أنَّ تعريف علم الإحصاء جاء لينسجم مع ما ذكره الله تعالى في تلك الآيات الكريمة، وأمّا من حيث شمولية تطبيق هذا العلم، فلا يكاد يخلو قسم من العلوم التطبيقية أو النظرية من دراسة الإحصاء، حيث تتلاقى في هذا النوع من العلوم نتائج من التحليل الرياضيّاتي وجبر المنطق والدراسات الاحتماليّة لتصبّ جميعاً في تحليل واستقراء مسألة أو ظاهرة يلفُّها الشّك والارتياب. لقد نُشر أول كتاب في الإحصاء عام 1845 من قبل الرّياضيّاتي (**الاكتواريّ**) الإنكليزي نيسون Francis Gustavus Paulus Neison وحمل عنوان **مساهمات في الإحصاء الحيويّ** "Contributions to Vital Statistics".

لقد نَشطَ في مجال الإحصاء عددٌ كبيرٌ جداً من المهتّمين بهذا العلم وأكثرهم من علماء العلوم الطّبيعيّة (الرّياضيات، الإحصاء، الفيزياء وعلم الحياة)، وكان من أبرّز هؤلاء: **بييز** Thomas Bayes (1702-1761)، **بواسون** Siméon Denis Poisson (1781-1840)، **بيرسون** Karl Pearson (1857-1936) و **فيشر** Ronald Fisher (1890-1962) **وآخرون** كثر...

أمّا **علم الاحتمالات** فإنّه يهتّم بدراسة التجارب والظواهر العشوائية ونمذجتها رياضياتياً، وهو الأخ الشقيق الأصغر للإحصاء من عائلة العشوائيّات Stochastic (الإحصاء الرّياضيّاتي ونظرية الاحتمالات)، وتاريخ ولادته حديث (نسبياً) إذا ما قُورن بتاريخ ولادة الإحصاء، فعلى الرغم من انتشاره الواسع جداً في أيامنا، فإنَّ مخاض ولادته كان عسيراً، فقد ولد من رحم دور الميسر وحلبات المراهنات، وقد عانى ما عناه من النقد حيناً والرفض كعلم حيناً آخر، فمنهم من رفض قبول الاحتمالات كعلم رياضياتي (كالرياضياتي برتراند Joseph Bertrand (1822 –1900))، ومنهم من صنّف الاحتمالات كفرع من الفيزياء (كالرياضياتي هيلبرت David Hilbert (1862 –1943)).

في الحقيقة نشأ الحساب الاحتمالي من خلال محاولة إيجاد الحلول لبعض ألعاب الحظ، وذكرت كتب التاريخ أنَّ حسابات لبعض ألعاب الحظ (لعبة حجر النرد) نُقلت عبر الحملات الصليبية من المنطقة العربية (من بلاد الشام) التي كانوا يسيطرون عليها، وذلك في القرن الثالث عشر الميلادي، وهذا أقدم ما دوِّن في هذا المجال، وذلك أنّه لم يُعثر في كتب التاريخ القديمة (كالهندية والصينية) ما يدلّ على وجود حسابات لألعاب الحظ، وبعد هذا النقل وجدَت حسابات فرديّة لألعاب الحظ لدى فئة كبيرة من الإيطاليين، ومن أشهرهم **كاردانو** Gerolamo Cardano (1501-1576). إنَّ القرن السابع عشر حمل في ثناياه تطوراً أكبر وأكثر دقّة لحسابات ألعاب الحظ على يد عدد من الرياضياتين الفرنسيين مثل: **باسكال** Blaise Pascal (1623-1662)، **فيرمات** Pierre de Fermat (1607-1665) **و هويغنز** Christiaan Huygens (1629-1695) الذي نشر أول كتاب في الحساب الاحتمالي، وكان موسوماً بـ "**حسابات لألعاب الحظ**" وذلك عام 1657. إلاَّ أنَّ عام 1713 يُعدُّ التاريخ الحقيقي لولادة علم الاحتمالات بعد أن تمَّ نشر أول مقالة علميَّة متقنة للرياضيّاتي السويسري **يعقوب برنولي** Jakob Bernoulli (1655-1705)، حيث نُشرت تلك المقالة من قِبل أخيه **يوهان** Johann Bernoulli (1667-1748) (بعد ثمان سنوات من وفات يعقوب)، وعلى إثر ذلك حظيت الاحتمالات باهتمام عدد كبير من علماء العلوم الطبيعيّة عامّة، وعلماء الرياضيات خاصّة لما لها من تطبيقات مهمّة في معظم مجالات العلوم التطبيقيّة والنظريّة على حد سواء. إنَّ النظرية الحديثة للاحتمالات بدأت من خلال ما يُعرف بـ "**مسلّمات كلموغوراف**" للفضاء الاحتمالي التي وضعت عام 1933 من قبل الرياضياتي الروسي (الاتحاد السوفياتي سابقا) **كلموغوراف** Andrey Kolmogorov (1903-1987). بعد ذلك تطوَّرت نظرية الاحتمالات وأصبحت بمثابة الحجر الأساس لمعظم الدراسات المهتمّة بالتجارب والظواهر العشوائية ونمذجتها.

بما يخصّ محتوى الكتاب، فقد قدَّمنا في طياته ثلاثة عشر فصلاً، وثلاثة ملاحق، منها ملحقين في الرياضيات، والثالث للجداول الإحصائية والخصائص العددية المميّزة لبعض المتغيِّرات العشوائية التي ذُكرت في الكتاب، وسرد للمراجع، وثبت للمصطلحات العلميّة، ودليل للأعلام الذين وردت أسماؤهم في هذا الكتاب. لقد تناولنا من خلال الفصول المقدَّمة موضوعات أساس وتطبيقات في الاحتمالات والإحصاء، وقد وزِّعت على ثلاثة محاور رئيسة هي:

١ – **الاحتمالات**، وقد خصِّص لها الفصول: **الرابع، والخامس، والسادس، والسابع والثامن**.

٢ – **الإحصاء**، وقد وزِّع على محورين اثنين هما:

أ– **الإحصاء الوصفي**، وقد خصِّص له الفصلين: **الأول والثاني**.

ب– **الإحصاء الاستدلالي**، وقد خصِّص له الفصول: **التاسع، والعاشر والحادي عشر**.

٣– **موضوعات متعلّقة بالاحتمالات والإحصاء**، وقد خصِّص لها الفصول: **الثالث، والثاني عشر، والثالث عشر**.

إنَّ توزيعنا لفصول هذا الكتاب على النّحو المذكور آنفًا يعود إلى سببين أساسيين هما:

١- **التدرج المعرفي للمعلومات المقدَّمة**: بمعنى أنَّنا أخذنا بالحسبان المستوى العلمي والمعرفي للقارئ، ولذلك بدأنا بمعلومات بسيطة متوفر بعضها لدى القارئ من مرحلة التعليم قبل الجامعي، ومن ثمَّ الولوج في الموضوعات الأكثر تجريداً وصعوبةً.

٢- **الترابط العلمي والموضوعي للفقرات المقدَّمة**: فقد حاولنا جاهدين ألاّ تظهر ثغرات بين المعلومات، بمعنى أنَّ المعلومات المتراكبة مبنيَّة على نحو متصّاعد، فإذا كان لدينا مفهوم يعتمد على آخر، فقد قمنا (وفي معظم الحالات) بذكر هذا الآخر أولاً ومن ثمَّ عرض المفهوم المعني.

في الحقيقة بذلنا جهداً كبيراً لتوضيح الفِكَر المقدَّمة بوساطة الأمثلة حيناً والجداول والعروض البيانية حيناً آخر، وعلاوة على ذلك قدمنا بعض التطبيقات باستخدام بعضِ البرامج الإحصائية وعلى وجه الخصوص برنامج Minitab مستخدمين في ذلك الإصدار السّابع عشر.

لقد زخرت فقرات هذا الكتاب وموضوعاته بعدد من الأمثلة المحلولة والتمارين غير المحلولة المناسبة للفِكَر المُقدَّمة في طياته، وكذلك رقّمت الفقرات، والعلاقات، والجداول، والأشكال بطريقة تُسهِّل على القارئ التنقل فيما بينها. كما قمنا باستخدام بعض معاجم الرياضيات من أجل عمليّة التعريب، وحاولنا جاهدين إغناء موضوعات هذا الكتاب بالقدر المستطاع من المعارف العامّة التي تناسب مرحلة الدراسة الجامعية الأولى لمختلف الاختصاصات، ولذلك تجنبنا التخصيص بالقدر المستطاع، وعرضنا النصوص والصيغ بشكل يمكن معه تحويرها ليناسب اختصاص القارئ. هذا بالنسبة لغير الاختصاصيين في الإحصاء والاحتمالات، وأمّا بالنسبة للمختصين في الإحصاء والاحتمالات فإنَّ محتويات هذا الكتاب تناسب (في الحد الأدنى) السنتين الدراسيتين الأوليين من مرحلة الدراسة الجامعية لذوي هذا الاختصاص.

إنَّ ما قُدِّم في طيات هذا الكتاب ليس إلاّ غيض من فيض، فكلّ موضوع من الموضوعات المقدَّمة يمكن تأليف كتب عنها، ولذلك قد يَجد القارئ الكثير من الاقتراحات على إضافة فقرة هنا، وحذف أخرى هناك، وتعديل على فقرة في موضع آخر، وهذا أمر قلّما يَسلم منه مُؤلِّف مهما بُذل من جهد لأجله، وذلك لأنّه ما زال هناك الكثير، والكثير جداً، من المعلومات التي يمكن أن تضاف في طيات الفصول السّابقة، ولكن لكل عمل نهاية، فمن وجد في كتاباتنا نقصاً فإنَّ الكمال لله تعالى وحده، ومن وجد خطأً فجلَّ من لا يخطئ. لذلك نودّ أن نعرب عن شكرنا وتقديرنا العميقين لكلّ من يبدي لنا ملاحظة مفيدة أو نقداً بنّاءً حول هذا الكتاب آملين من اللهّ تعالى أن نكون قد وفقنا في تقديمه بالشكل المناسب والمفيد.

**في الختام يودّ المؤلِّفان تقديم الشكر الجزيل إلى عمادة البحث العلمي في جامعة الملك سعود التي رعت دعم تأليف هذا الكتاب، كما يشكر المؤلِّفان جامعة الملك سعود التي أسهمت في طباعة هذا الكتاب ونشره.**

**اللَّهم انفعنا بما علمتنا وعلّمنا ما ينفعنا وزدنا علماً وعملاً متقبلاً إنَّك أنت السميع العليم**

**الرياض: يوم الأربعاء في ١٧/ ١١/ ١٤٣٨ الموافق لـ ٠٩/ ٠٨/ ٢٠١٧**

**المؤلِّفان**

**أ.د. حميد عويِّد مشرف العكله** haloklah@ksu.edu.sa

**د. عبيد جفين جليغم القحطاني** obalgahtani@ksu.edu.sa

# المحتويات


**الفصل الأول: البيانات الإحصائية وطرائق عرضها** ........ ١

١, ١ تعاريف ومفاهيم أساسية ........ ١

١, ١, ١ علم الإحصاء ........ ١

١, ١, ٢ المجتمع الإحصائي ........ ١

١, ١, ٣ العيِّنة ........ ٢

١, ١, ٤ البيانات الإحصائيّة ........ ٣

١, ٢ المتغيِّرات وأنواع البيانات ........ ٥

١, ٢, ١. المتغيِّرات ........ ٥

١, ٢, ٢ البيانات النوعيّة ........ ٦

١, ٢, ٣ البيانات الكميَّة ........ ٩

١, ٣ تمثيل البيانات وعرضها ........ ١١

١, ٣, ١ طريقة الجدول ........ ١٣

١, ٣, ٢ التمثيل باستخدام لوحة الانتشار ........ ١٦

١, ٣, ٣ العرض الشرائطي ........ ١٧

١, ٣, ٤ المضلَّع ........ ١٩

١, ٣, ٥ الخط المنحني ........ ٢١

١, ٣, ٦ القرص الدائري ........ ٢١

١, ٣, ٧ التمثيل بالصور ........ ٢٣

١, ٤ الجداول التكرارية ذوات الفئات ........ ٢٤

١, ٤, ١ الفئة ........ ٢٤

١ , ٤ , ٢ سعة الفئة ........ ٢٤
١ , ٤ , ٣ تكرار الفئة ........ ٢٤
١ , ٤ , ٤ الجداول التكرارية ذوات الفئات المتّصلة ........ ٢٦
١ , ٤ , ٥ الجداول التكرارية ذوات الفئات المتّصلة ........ ٢٦
١ , ٤ , ٦ حساب سعات الفئات وعددها من أجل جدول تكراري ذي فئات ........ ٢٧
١ , ٤ , ٧ جداول التوزيع التكرارية ........ ٣٠
١ , ٥ العروض البيانية لبيانات جداول التوزيع التكرارية ........ ٣٦
١ , ٥ , ١ المدرَّجات التكرارية ........ ٣٦
١ , ٥ , ٢ المضلَّعات التكرارية ........ ٤١
١ , ٥ , ٣ المضلَّعات التكرارية التراكمية ........ ٤٢
١ , ٥ , ٤ المنحنيات التكرارية ........ ٤٣
١ , ٦ أشكال التوزيعات التكرارية ........ ٤٨
١ , ٦ , ١ التوزيعات التكرارية المتناظرة ........ ٤٨
١ , ٦ , ٢ التوزيعات التكرارية الملتوية ........ ٤٩
١ , ٦ , ٣ تفسير شكل التوزيع ........ ٥٠
تمارين الفصل الأول ........ ٥١

الفصل الثاني: المقاييس الوصفية للبيانات ........ ٥٥
٢ , ١ مقاييس النزعة المركزية ........ ٥٦
٢ , ١ , ١ تعريف مقياس النزعة المركزية ........ ٥٦
٢ , ١ , ٢ المتوسط ........ ٥٦
٢ , ١ , ٣ المتوسط الموزون ........ ٦١
٢ , ١ , ٤ المتوسط الهندسي ........ ٦٢
٢ , ١ , ٥ المتوسط التوافقي ........ ٦٤
٢ , ١ , ٦ الوسيط ........ ٦٦
٢ , ١ , ٧ المنـوال ........ ٧٠
٢ , ١ , ٨ مقارنة بين صفات المتوسط، والوسيط، والمنوال ........ ٧٢

٢,٢ الرُّبَيعيات ........ ٧٣
١,٢,٢ تعريف الربيعيات ........ ٧٣
٢,٢,٢ تعيين الربيعيات ........ ٧٤
٣,٢ المئينات ........ ٧٧
١,٣,٢ تعريف المئيني الرائيّ (ذو الرقم r) ........ ٧٧
٢,٣,٢ تعيين المئينات ........ ٧٨
٤,٢ مقاييس التشتت ........ ٨٢
١,٤,٢ الانحراف المتوسط ........ ٨٣
٢,٤,٢ التباين ........ ٨٥
٣,٤,٢ قاعدة تشبيشيف التجريبية ........ ٨٩
٤,٤,٢ المدى ........ ٩٠
٥,٤,٢ المدى الربيعي ........ ٩١
٦,٤,٢ معامل التَغيُّر ........ ٩٢
٧,٤,٢ معامل التشتت ........ ٩٣
٨,٤,٢ الدرجة المعيارية $Z$ ........ ٩٤
٩,٤,٢ الدرجة المعيارية $T$ ........ ٩٥
٥,٢ مقاييس الشكل للتوزيعات التكرارية ........ ٩٦
١,٥,٢ العزوم المركزية لمجموعة بيانات ........ ٩٦
٢,٥,٢ مقاييس الالتواء ........ ٩٩
٣,٥,٢ مقاييس التفلطح ........ ١٠٥
٦,٢ الأعداد الخمسة والتمثيل الصندوقي للبيانات ........ ١٠٧
١,٦,٢ الأعداد الخمسة لبيانات كمّيّة خام ........ ٩٦
٢,٦,٢ التمثيل الصندوقي لبيانات كمّيّة خام ........ ٩٩
تمارين الفصل الثاني ........ ١١١

الفصل الثالث: الارتباط وتحليل الانحدار ........ ١١٥
١,٣ الارتباط الخطي البسيط ........ ١١٥

١ ,١ ,٣ نماذج من الارتباط ........ ١١٥
٢ ,١ ,٣ لوحة الانتشار ........ ١١٩
٣ ,١ ,٣ التغاير لمجموعتي بيانات ........ ١٢٢
٤ ,١ ,٣ مُعامل بيرسون للارتباط ........ ١٢٧
٥ ,١ ,٣ مُعامل سبيرمان لارتباط الرتب ........ ١٣٠
٦ ,١ ,٣ مُعامل الاقتران ........ ١٤٠
٧ ,١ ,٣ مُعامل التوافق ........ ١٤١
٢ ,٣ الانحدار الخطي البسيط ........ ١٤٢
١ ,٢ ,٣ طريقة المربعات الصغرى ........ ١٤٤
٢ ,٢ ,٣ معادلتا مستقيما الانحدار لبيانات عينة ........ ١٤٥
٣ ,٢ ,٣ معادلتا مستقيما الانحدار لبيانات مجتمع إحصائي ........ ١٥٤
٤ ,٢ ,٣ دقة التقديرات ........ ١٥٨
٣ ,٣ الانحدار غير الخطي وتوفيق المنحنيات ........ ١٦٥
١ ,٣ ,٣ نماذج انحدار يمكن ردّها إلى خطية ........ ١٦٧
٢ ,٣ ,٣ نماذج لا يمكن خطيتها ........ ١٧٠
٤ ,٣ الارتباط المتعدِّد والجزئي ........ ١٧٠
١ ,٤ ,٣ الارتباط المتعدِّد ........ ١٧٠
٢ ,٤ ,٣ العلاقة بين الارتباط المتعدِّد والارتباط البسيط ........ ١٧٥
٣ ,٤ ,٣ الارتباط الجزئي ........ ١٧٨
٤ ,٤ ,٣ العلاقة بين الارتباط المتعدِّد والارتباط الجزئي ........ ١٨١
تمارين الفصل الثالث ........ ١٨٣

الفصل الرابع: الفضاء الاحتمالي، الاحتمالات الشرطية واستقلال الحوادث ........ ١٨٧
١ ,٤ الفضاء الاحتمالي لتجربة عشوائية ........ ١٨٧
١ ,١ ,٤ تعيين مجموعة نتائج تجربة عشوائية ........ ١٨٩
٢ ,١ ,٤ جبر الحوادث ........ ١٩٥
٣ ,١ ,٤ تعيين الدالَّة الاحتمالية ........ ٢٠٤

٤ , ١ , ٤ خصائص الدوال الاحتمالية ........ ٢١٤
٤ , ٢ الاحتمالات الشرطية ........ ٢١٨
١ , ٢ , ٤ الاحتمال الشرطي لحادث ........ ٢١٨
٢ , ٢ , ٤ قانون الضرب في الاحتمالات ........ ٢٢٠
٣ , ٢ , ٤ صيغة الاحتمال التام ........ ٢٢٠
٤ , ٢ , ٤ نظرية بييز ........ ٢٢١
٣ , ٤ استقلال الحوادث ........ ٢٢٣
١ , ٣ , ٤ استقلال حادثين ........ ٢٢٣
٢ , ٣ , ٤ استقلال عدد منته من الحوادث ........ ٢٢٤
٣ , ٣ , ٤ خصائص الحوادث المستقلّة ........ ٢٢٦
تمارين الفصل الرابع ........ ٢٢٩

الفصل الخامس: المتغيِّرات العشوائية وتوزيعاتها الاحتمالية ........ ٢٣٣
١ , ٥ المتغيِّرات العشوائية ........ ٢٣٣
١ , ١ , ٥ تعريف المتغيِّر العشوائي ........ ٢٣٤
٢ , ١ , ٥ تعريف العنصر العشوائي ........ ٢٣٨
٣ , ١ , ٥ الدالَّة المميِّزة لحادث ........ ٢٣٩
٤ , ١ , ٥ الدوال في متغيّرات عشوائية ........ ٢٣٩
٢ , ٥ قانون ودالَّة توزيع متغيِّر عشوائي ........ ٢٤٠
١ , ٢ , ٥ قانون توزيع متغيِّر عشوائي ........ ٢٤٠
٢ , ٢ , ٥ دالَّة التوزيع لمتغيِّر عشوائي ........ ٢٤٣
٣ , ٢ , ٥ الخصائص المميِّزة لدوال التوزيع الاحتمالية على R ........ ٢٤٤
٤ , ٢ , ٥ الشرط اللازم والكافي لاستمرار دالَّة توزيع احتمالية على R ........ ٢٤٨
٣ , ٥ تصنيف المتغيِّرات العشوائية ........ ٢٥١
١ , ٣ , ٥ المتغيِّرات العشوائية المتقطِّعة ........ ٢٥٢
٢ , ٣ , ٥ المتغيّرات العشوائية المستمرّة ........ ٢٥٨
٤ , ٥ المتَّجهات العشوائيَّة وتوزيعاتها الاحتمالية ........ ٢٦٥

٥ , ٤ , ١ مفهوم المتّـجه العشوائي ........................................ ٢٦٦
٥ , ٤ , ٢ دالّة التوزيع الاحتمالية لمتّـجه عشوائي ........................................ ٢٦٦
٥ , ٤ , ٣ خصائص دوال التوزيع الاحتمالية على $\mathbb{R}^n$ ........................................ ٢٦٨
٥ , ٤ , ٤ تصنيف المتّجهات العشوائيّة ........................................ ٢٧٢
٥ , ٥ التوزيعات الهامشية ........................................ ٢٧٨
٥ , ٥ , ١ دوال الكثافة الهامشية ........................................ ٢٧٨
٥ , ٥ , ٢ دوال التوزيع الهامشية ........................................ ٢٧٩
٥ , ٦ التوزيعات الشرطية ........................................ ٢٨٣
٥ , ٦ , ١ الاحتمالات الشرطية للمتغيّرات العشوائي ........................................ ٢٨٣
٥ , ٦ , ٢ التوزيعات الشرطية لمتّجه عشوائي متقطّع ........................................ ٢٨٤
٥ , ٦ , ٣ التوزيعات الشرطية لمتّجه عشوائي مستمر ........................................ ٢٨٩
تمارين الفصل الخامس ........................................ ٢٩٦

الفصل السادس: توزيعات احتمالية شهيرة ........................................ ٣٠١
٦ , ١ توزيعات احتمالية متقطّعة ........................................ ٣٠١
٦ , ١ , ١ التوزيع الوحيد النقطة ........................................ ٣٠١
٦ , ١ , ٢ التوزيع الثنائي النقطة ........................................ ٣٠٣
٦ , ١ , ٣ التوزيع المنتظم المتقطّع ........................................ ٣٠٧
٦ , ١ , ٤ التوزيع الحدّاني ........................................ ٣٠٩
٦ , ١ , ٥ التوزيع الحدّاني السالب ........................................ ٣١١
٦ , ١ , ٦ التوزيع الهندسي ........................................ ٣١٤
٦ , ١ , ٧ التوزيع فوق الهندسي ........................................ ٣١٨
٦ , ١ , ٨ توزيع بواسون ........................................ ٣٢٠
٦ , ٢ توزيعات احتمالية مستمرّة ........................................ ٣٢٣
٦ , ٢ , ١ التوزيع المنتظم المستمر ........................................ ٣٢٣
٦ , ٢ , ٢ التوزيع الأسّي ........................................ ٣٢٤
٦ , ٢ , ٣ التوزيع الطبيعي ........................................ ٣٢٧

٦,٢,٤ توزيع ستودنت ........ ٣٣٥
٦,٢,٥ توزيع كاي مربَّع ........ ٣٣٨
٦,٢,٦ توزيع فيشر ........ ٣٤٠
٦,٣ تقارب بعض التوزيعات الاحتمالية من بعضها الآخر ........ ٣٤٢
٦,٣,١ مبرهنة (بخصوص تقارب التوزيع فوق الهندسي من الحدَّاني) ........ ٣٤٢
٦,٣,٢ مبرهنة (بخصوص تقارب الحدَّاني من البواسوني) ........ ٣٤٣
٦,٣,٣ مبرهنة (بخصوص تقارب الحدَّاني السالب من البواسوني) ........ ٣٤٣
تمارين الفصل السادس ........ ٢٤٧

الفصل السابع: استقلال وعزوم المتغيِّرات العشوائية ........ ٣٥١
٧,١ استقلال المتغيِّرات العشوائية ........ ٣٥١
٧,١,١ الاستقلال لعدد منته من المتغيِّرات العشوائية ........ ٣٥١
٧,١,٢ الاستقلال لعدد غير منته من المتغيّرات العشوائية ........ ٣٥٦
٧,٢ العزوم للمتغيِّرات العشوائية ........ ٣٥٨
٧,٢,١ العزم الابتدائي من المرتبة k لمتغيِّر عشوائي ........ ٣٥٨
٧,٢,٢ العزوم المركزية لمتغيِّر عشوائي ........ ٢٧٢
٧,٢,٣ العزم العامليّ (الضربيّ) لمتغيِّر عشوائي ........ ٣٧٦
٧,٢,٤ حساب العزوم لبعض المتغيِّرات العشوائية الشهيرة ........ ٣٧٦
٧,٢,٥ متباينات شهيرة ذات صلة بالعزوم ........ ٣٨١
٧,٢,٦ مُعامل الارتباط الخطي لمتغيِّرين عشوائيين ........ ٣٨٢
٧,٣ الدوال المولِّدة الاحتمالية ........ ٣٨٣
٧,٣,١ تعريف الدالَّة المولِّدة الاحتماليةلتغيِّر عشوائي ........ ٣٨٣
٧,٣,٢ خصائص الدالَّة المولِّدة الاحتمالية ........ ٣٨٤
٧,٣,٣ الدالَّة المولِّدة لعزوم متغيِّر عشوائي ........ ٣٨٩
٧,٤ الدوال المميِّزة للتوزيعات الاحتمالية ........ ٣٩٢
٧,٤,١ الدالَّة المميِّزة لتوزيع احتمالي ........ ٣٩٢
٧,٤,٢ خصائص الدالَّة المميِّزة لتوزيع احتمالي ........ ٣٩٣

٣,٤,٧ نظريات شهيرة في دراسة الدوال المميِّزة ........ ٣٩٦
٥,٧ التوقع الشرطي ........ ٣٩٨
١,٥,٧ التوقع الشرطي لحادث A بالنسبة إلى تجزئة ........ ٣٩٨
٢,٥,٧ التوقع الشرطي لمتغيِّر عشوائي بالنسبة إلى تجزئة ........ ٤٠٠
تمارين الفصل السابع ........ ٤٠٧

الفصل الثامن: بعض نظريات النهاية المركزية ........ ٤١١
١,٨ تقارب متتاليات المتغيِّرات العشوائيَّة ........ ٤١٢
١,١,٨ أشهر أنواع التقارب لمتتاليات المتغيِّرات العشوائيَّة ........ ٤١٢
٢,١,٨ تقارب متتاليات المتغيِّرات العشوائية ........ ٤١٣
٢,٨ قانونا الأعداد الكبيرة ........ ٤١٤
١,٣,٨ تعريف قانوني الأعداد الكبيرة (الضعيف والقوي) ........ ٤١٤
٢,٣,٨ مبرهنة بخصوص تقارب المتتالية $\left(\frac{S_n}{n}\right)_{n\in\mathbb{N}}$ ........ ٤١٥
٣,٣,٨ مبرهنة بخصوص تقارب المتتالية $\left(\frac{S_n}{n}\right)_{n\in\mathbb{N}}$ من ثابت حقيقي $\mu$ ........ ٤١٦
٣,٨ نظريات النهاية المركزية ........ ٤١٧
١,٣,٨ مسألة موافِر- لابلاس ........ ٤١٧
٢,٣,٨ مسألة ليندنبرغ – ليفي ........ ٤٢١
٣,٣,٨ مسألة ليابانوف ........ ٤٢٨
٤,٣,٨ مسألة ليندنبرغ- فيلر ........ ٤٣١
٥,٣,٨ مسألة غنيدينكو ........ ٤٣٤
تمارين الفصل الثامن ........ ٤٣٧

الفصل التاسع: مدخل إلى نظرية المعاينة ........ ٤٣٩
١,٩ مفاهيم ومصطلحات أساسية ........ ٤٣٩
٢,٩ تصنيف العيِّنات ........ ٤٤١
١,٢,٩ العيِّنات العشوائية ........ ٤٤١

٩,٢,٢ العيِّنات العمديَّة ........ ٤٤٧

٩,٣ البناء الرياضياتي النظري للعيِّنة العشوائية البسيطة ........ ٤٤٨

٩,٤ النموذج الرياضياتي النظري لمجتمع إحصائي ........ ٤٥٠

٩,٥ الإحصاءات ودوال توزيع العينات ........ ٤٥١

٩,٥,١ الإحصاءة ........ ٤٥١

٩,٥,٢ دالَّة التوزيع العمليَّة للعيِّنة ........ ٤٥٣

٩,٦ المميِّزات العددية لبعض الإحصاءات ........ ٤٥٤

٩,٧ توزيعات المعاينة في حال مجتمع وحيد ........ ٤٥٦

٩,٧,١ توزيعات المعاينة لمتوسط وتباين عينة من مجتمع طبيعي ........ ٤٥٦

٩,٧,٢ توزيع المعاينة لمتوسط (لنسبة) عينة من مجتمع برنولي ........ ٤٥٩

٩,٨ توزيعات المعاينة في حال مجتمعين مستقلَّين ........ ٤٦١

٩,٨,١ توزيعات المعاينة لفرق متوسطي عينتين ونسبة التباين لعينتين من مجتمعين طبيعيين مستقلَّين ........ ٤٦٢

٩,٨,٢ توزيعات المعاينة للفرق بين متوسطي (نسبتي) عينتين من مجتمعين برنوليين مستقلَّين ........ ٤٦٦

٩,٩ الإحصاءة الكافية ........ ٤٦٧

٩,١٠ كمية معلومات فيشر التي تُقدِّمها عيِّنة عشوائية ........ ٤٧١

تمارين الفصل التاسع ........ ٤٧٥

الفصل العاشر: مدخل إلى نظرية التقدير ........ ٤٧٧

١٠,١ تعيين مُقدِّر مَعْلمة مجتمع إحصائي ........ ٤٧٨

١٠,١,١ طريقة العزوم في تعيين مُقدِّرات مَعَالم مجتمع ........ ٤٧٨

١٠,١,٢ طريقة الأرجحية العظمى في تعيين مُقدِّرات مَعَالم مجتمع ........ ٤٨٠

١٠,٢ التقدير النقطي ........ ٤٨٥

١٠,٢,١ المقدِّرات غير المنحازة (أو المُنْصِفة) ........ ٤٨٦

١٠,٢,٢ المقدِّرات المتَّسقة ........ ٤٨٨

١٠,٢,٣ المقدِّرات الكفؤة ........ ٤٨٩

١٠,٢,٤ متباينة راو كرامر ........ ٤٩٠

١٠,٣ فترات الثقة ........ ٤٩٢

١, ٣, ١٠ مفهوم فترة الثقة لَمْعلمة مجتمع إحصائي ........ ٤٩٢
٢, ٣, ١٠ فترات الثقة لَمْعلمتي مجتمع طبيعي ........ ٤٩٤
٣, ٣, ١٠ فترة الثقة لَمْعلمة مجتمع برنولي ........ ٤٩٨
٤, ٣, ١٠ فترات الثّقة للفرق بين متوسطي مجتمعين طبيعيين مستقلَّين ولنسبة تشتتهما ........ ٥٠١
٥, ٣, ١٠ فترة الثقة للفرق بين مَعْلمتي مجتمعين برنوليين مستقلَّين ........ ٥٠٤
تمارين الفصل العاشر ........ ٥١٠

الفصل الحادي عشر: مدخل إلى اختبارات الفرضيات الإحصائية ........ ٥١٣
١, ١١ مفاهيم أساس في دراسة اختبار الفرضيات الإحصائية ........ ٥١٣
١, ١, ١١ الفرضيات الإحصائية ........ ٥١٣
٢, ١, ١١ الأخطاء الناتجة عن دراسة فرضية إحصائية واحتمالاتها ........ ٥١٤
٣, ١, ١١ تعيين منطقة الرفض المثلى للفرضية الابتدائية ........ ٥١٩
٢, ١١ الاختبارات الإحصائية من أجل معلمتي مجتمع طبيعي ........ ٥٢١
١, ٢, ١١ اختبارات الإحصائية من أجل متوسط مجتمع طبيعي ........ ٥٢١
٢, ٢, ١١ اختبارات من أجل تباين مجتمع طبيعي ........ ٥٢٦
٣, ١١ الاختبارات الإحصائية من أجل مَعْلَمة مجتمع برنولي ........ ٥٢٩
٤, ١١ اختبارات من أجل الفرق بين متوسطي مجتمعين طبيعيين مستقلَّين ........ ٥٣٠
٥, ١١ اختبارات من أجل نسبة التباين لمجتمعين طبيعيين مستقلَّين ........ ٥٣٢
٦, ١١ اختبار من أجل الفرق بين مَعْلَمتي مجتمعين برنوليين مستقلَّين ........ ٥٣٥
٧, ١١ اختبارات إحصائية من أجل فرضيات غير مَعْلَميّة ........ ٥٣٦
١, ٧, ١١ اختبارات جودة توفيق التوزيعات الاحتمالية ........ ٥٣٧
٢, ٧, ١١ اختبارات إحصائية من أجل تجانس التوزيعات ........ ٥٤٤
٣, ٧, ١١ اختبار استقلال ظاهرتين (أو متغيِّرين) ........ ٥٤٨
٨, ١١ تطبيقات ........ ٥٥١
تمارين الفصل الحادي عشر ........ ٥٥٩

**الفصل الثاني عشر: المتسلسلات الزمنية ........ ٥٦٣**
١٢,١ مفاهيم أولية في المتسلسلات الزمنية ........ ٥٦٣
١٢,١,١ المتسلسلة الزمنية ........ ٥٦٣
١٢,١,٢ الطوري العشوائي ........ ٥٦٤
١٢,١,٣ بصمة متسلسلة زمنية ........ ٥٦٥
١٢,١,٤ مُركّبات المتسلسلة الزمنية ........ ٥٦٩
١٢,٢ طريقة التجزئة في تحليل متسلسلة زمنية ........ ٥٧٢
١٢,٣ تحليل الاتجاه العام لمتسلسلة زمنية ........ ٥٧٦
١٢,٣,١ تقدير الاتجاه العام بوساطة مستقيم ........ ٥٧٦
١٢,٣,٢ تقدير الاتجاه العام بوساطة كثيرة حدود ........ ٥٧٨
١٢,٣,٣ تقدير الاتجاه العام بوساطة دوال النمو ........ ٥٨٣
١٢,٣,٤ اختبار الفرق لـ نويمان ........ ٥٨٧
١٢,٤ دراسة المُركّبة الدوريّة لمتسلسلة زمنية ........ ٥٩١
١٢,٤,١ طريقة جيب التمام (أو التجيب) لـ هالبرغ ........ ٥٩١
١٢,٤,٢ اختبار جودة التوفيق للمُركّبة الدوريّة المقدّرة بوساطة طريقة جيب التمام لـ هالبرغ ........ ٥٩٧
١٢,٥ تنعيم المتسلسلات الزمنية ........ ٦٠١
١٢,٥,١ مُعامل الخشونة لمتسلسلة زمنية ........ ٦٠١
١٢,٥,٢ التنعيم باستخدام المتوسطات المتحركة ........ ٦٠٢
١٢,٦ دراسة التغيُّرات الموسميّة لمتسلسلة زمنية ........ ٦٠٤
١٢,٦,١ طريقة النسبة إلى متوسط متحرك ........ ٦٠٥
١٢,٧ طريقة وينتر لتوليد تنبؤات من متسلسلة زمنية ........ ٦٠٨
١٢,٧,١ طريقة وينتر في توفيق نموذج متسلسلة زمنية ........ ٦٠٩
تمارين الفصل الثاني عشر ........ ٦١٢

**الفصل الثالث عشر: الأرقام القياسية ........ ٦١٥**
١٣,١ مفهوم الرقم القياسي ........ ٦١٥
١٣,١,١ الرقم القياسي ........ ٦١٨

١٣,٢ الأرقام القياسية البسيطة ........ ٦١٨

١٣,٢,١ الرقم القياسي البسيط ........ ٦١٨

١٣,٢,٢ الرقم القياسي الكلي البسيط ........ ٦٢٣

١٣,٣ الأرقام القياسية الموزونة ........ ٦٢٥

١٣,٣,١ مؤشِّر لسبيريس ........ ٦٢٥

١٣,٣,٢ مؤشِّر باشّي ........ ٦٢٨

١٣,٣,٣ المؤشِّر المثالي لـ فيشر ........ ٦٣١

١٣,٤ أمثلة ........ ٦٣٢

تمارين الفصل الثالث عشر ........ ٦٣٧

الملحق A ........ ٦٤١

الملحق B ........ ٦٥٣

الملحق C ........ ٦٦٧

ملخَّص لتوزيعات احتمالية شهيرة ........ ٦٨٩

المراجع ........ ٦٩٥

دليل الأعلام ........ ٦٩٩

ثبت المصطلحات

عربي ⇐ English ........ ٧٠٣

عربي ⇒ English ........ ٧٢٠

# الفصل الأول

## البيانات الإحصائية وطرائق عرضها
## STATISTICAL DATA AND METHODS OF THEIR PRESENTATION

### (١,١) تعاريف ومفاهيم أساسية
### Definitions and Basic Concepts

**بادئ ذي بدء نشير إلى أنَّ جميع رموز مجموعات الأعداد التي سترد في هذا الكتاب، وكذلك رموز مُدْرَجة وأخرى رياضياتية، تجدها في بداية الملحق (A).**

لقد تحدثنا في مقدِّمة الكتاب عن تاريخ علم الإحصاء، وذكرنا أنَّه علم قديم نسبيًّا وله مدلول واضح عند الكثير من النّاس، ولكن وفقاً لمفاهيم العدّ والتصنيف. أمّا مدلولاته وفقاً للدراسات العلميّة فإنَّه يحتاج إلى صياغات علميّة دقيقة، وهذا ما سنقوم به في كتابنا هذا.

نقدِّم فيما يلي بعض التعاريف التي لا غنى عنها لدى البدء في دراسة الإحصاء الوصفي، علماً أنَّ المراجع حول مواضيع هذا الفصل والفصل الذي يليه كثيرة جداً ومنها ما ذُكِر في مراجع هذا الكتاب، والاستدلال عليها واضح من عناوين تلك المراجع.

**(١ , ١ , ١) تعريف (علم الإحصاء Statistics)**

علم الإحصاء هو أحد فروع الرّياضيّات الذي يهتمّ بجمع البيانات ومعالجتها (**كالبحث في نوعها وكيفية تقديمها، وكذلك النظر في نواقصها والمفقود منها**)، وتمثيلها في جداول وأشكال مناسبة، واستنتاج بعض المقاييس العدديّة بشأنها، وتحليلها، ومن ثمَّ استقرائها، وأخيراً اتخاذ القرارات المناسبة بشأنها.

أمّا أهميّة هذا العلم فلم تعدْ تخفى على دارسي العلوم كافة، فيمكننا الجزم بأنَّه لم يُعد هناك فرع من فروع العلوم العلميّة (**كالفيزياء والكيمياء والطب والهندسات و...**) والإنسانيّة (**كعلم الاجتماع والآداب والسياسة و...**) إلاّ وقد ولج علم الإحصاء في بعض دراساته. في الواقع إنَّ الدراسة الإحصائية لأيّة مسألة تنطلق ممّا يُعرف باسم **المجتمع الإحصائي** الذي يكوِّن الركيزة الأساسية للبيانات التي ستخضع للدراسة، ولذلك لابدّ من تقديم تعريف واضحٍ لمعنى المجتمع الإحصائي كي لا يكون هناك أي لَبسٍ في ذهن القارئ لهذا المفهوم المهمّ.

**(٢ , ١ , ١) تعريف (المجتمع الإحصائي Population)**

إنَّ أي تجمُّع لأشياء تجمَع بينها صفة مشتركة واحدة على الأقل لتكون محل دراسة لهدف معيَّن تُدعى **مجتمعاً إحصائياً** (**أو مجتمعاً وذلك على سبيل الاختصار**)، وأمّا مكوّنات المجتمع الإحصائي فإنّها تُدعى **عناصر** أو **أفراداً**.

فيما يلي نقدّم بعض الأمثلة التي توضِّح هذا المفهوم:

١- في دراسة لتعيين متوسط أعداد الطلاب في المدارس الثّانوية في مدينة الرياض، حيث توجد مراحل مختلفة للطلاب مثل الأوّل، والثّاني، والثّالث الثّانوي، وكذلك يوجد الذكور والإناث، ولكنّهم جميعاً تجمع بينهم صفة الدّراسة في المرحلة الثانوية، ولذلك نجد أنَّ طلاب المرحلة الثانوية في الرياض يكوِّنون مجتمعاً إحصائياً.

٢- لدى دراسة تطور الحالة المرضيّة لمرضى في مستشفى ما، حيث يوجد أشخاص كُثر وبأنواع مختلفة من الأمراض، ولكنَّهم جميعاً تربط بينهم صفة الإصابة بمرض، ولذلك مرضى ذلك المستشفى يكوِّنون مجتمعاً إحصائياً.

٣- في دراسة تأثير مضاد حيويّ على جراثيم (بكتريا) في مزرعة جرثوميّة ما، حيث توجد أنواع عديدة منها العنقودية والعصيات و...، ولكن جميعها تندرج تحت اسم الجراثيم، ولذلك الجراثيم الموجودة في هذه المزرعة تكوِّن مجتمعاً إحصائياً.

٤- في دراسة تحديد درجة حلاوة التمور المنتجة من أشجار نخيل التّمر في المملكة العربيّة السعوديّة، حيث يوجد الكثير من أنواعه مثل: البرحي، المدينة، السكري و...، ولكنّـها جميعها تنضوي تحت صفة التّمر، ولذلك تمور تلك الأشجار تكوِّن مجتمعاً إحصائياً.

٥- في إطار الفحص الدوري للسّيّارات في مدينة ما، حيث توجدالسيّارات السّيّاحية والحافلات والشّاحنات و...، ولكن جميعها تعرف باسم السيارات لأنَّها تسير على الأرض بعجلاتها، ولذلك فإنَّ سيارات تلك المدينة تكوِّن مجتمعاً إحصائياً.

٦- لدى تحديد نسبة خام الحديد في فلز الحديد في منجم معيّن، حيث يمكن أن توجد أنواع عديدة من مُركَّبات الحديد في هـذا المنجم مثل: كبريتات الحديد والأكاسيد المختلفة للحديد و...، ولكنَّها جميعاً تحوي معدن الحديد؛ ولذلك فلز الحديـد في هـذا المـنجم يكوِّن مجتمعاً إحصائياً.

٧- في دراسة تحديد الحالة الفنية لطائرات دول مجلس التعاون الخليجي، حيث توجد فيها الطائرات السفريّة والحربيّة والتدريبيّة والحوامات و...، ولكنَّها جميعاً تتصف بقدرتها على الطيران، ولذلك الطائرات الموجودة في دول مجلس التعاون الخليجي تكوِّن مجتمعاً إحصائياً.

نلاحظ من هذه الأمثلة أنَّ المجتمع الإحصائي ليس بالضرورة أن يكون مجتمعاً بشرياً أو حتى مجتمعاً لأحياء، إذ إنَّه من الممكن أن يكون جماداً أيضاً.

**(١,٢,١,١) ملاحظة**

إنَّ عدد العناصر المكوِّنة للمجتمع يُدعى **حجم المجتمع** Size of Population، وفي حال كان حجم المجتمع منتهياً فإنَّه يُرمز لحجمه بحرف لاتيني كبير من قبيل $N$ أو $M$ أو...

### (١,١,٣) العيِّنة Sample

لقد ذكرنا أنَّ المجتمع الإحصائي يكوِّن الركيزة الأساسية للبيانات التي ستخضع للدراسة، ولكن قد تكون عمليّة إخضاع جميع عناصر المجتمع للبحث والدراسة (وتُدعى المسح الشامل **Comprehensive Survey**) شاقّة جداً، لا بل قد يكون في كثير من الحالات غير ممكن أيضاً، والمثالان الآتيان يوضِّحان لنا ذلك:

١ - بفرض أنَّ باحثاً يريد دراسة سلوك الأسماك البحريّة، فإنَّه لا يمكن لأي عاقل أن يدَّعي أنَّه قادر على مسح كافة الأسماك الموجودة في بحار الكرة الأرضيّة، ولو افترضنا جدلاً أنَّ ذلك ممكن، فتصور الكلفة الماديّة والزمنية المترتّبة على هذه الدراسة.

٢- بفرض أنَّ لجنة الرقابة على الأدوية تريد التحقُّق من مكوّنات عقار دوائيّ معيّن معبأ في كبسولات (وليكن على سبيل المثال

خافض الحرارة ومسكن الألم Setamol 500 mg )، فعندئذ من غير المجدي والمعقول أن تقوم هذه اللجنة بإخضاع كافة المنتج للتحليل المخبريّ (ومن ثمَّ إتلاف كافة الإنتاج ) للتثبّت من أنَّ المنتَج مُحقِّقاً للمواصفات المقدَّمة من الصانع.

إذن، والحال كذلك كان لا بدَّ من إيجاد وسيلة فعّالة نستطيع من خلالها الاستدلال على خصائص المجتمع بتكلفة مادية وزمنية معقولة، وفي الوقت نفسه تُجنّبنا بعض السلبيّات التي ذكرت آنفاً. في هذا الصدد بَدت فكرة التعامل مع جزء من المجتمع من أجل تنفيذ الدراسة الإحصائية. إنَّ العلم الذي يبحث في كيفية أخذ هذا الجزء من المجتمع واستخلاص البيانات المطلوبة للدراسة الإحصائية (ومن ثمَّ دراسة خصائصه ) يُعرف باسم **نظرية المعاينة** Sampling Theory. في الواقع تُعدُّ نظرية المعاينة جزءاً لا يتجزأ من العلوم الإحصائية المختلفة وخاصة بعد أن أصبح لهذه النظرية أسسٌ رياضياتية ثابتة ومكانة متميِّزة في المجالات التطبيقية المختلفة، كما أنَّ تطبيق هذه النظرية في البحوث العلمية تُمكننا من الحصول على نتائج سريعة وبكلفة زهيدة نسبياً إذا ما قُورنت بكلفة المسح الشامل. ليس هذا فحسب، لا بل إنَّه في كثير من الحالات يُنظر إلى هذه النظرية على أنَّها الوسيلة الوحيدة في البحَث العلمي وذلك للأسباب التي ذكرناها فيما سبق. إنَّ الفقرة الآتية تقدِّم لنا تعريف العيِّنة الذي يُعدّ بمثابة حجر الأساس لهذه النظريّة.

### (١ , ٣ , ١ , ١) تعريف (العيِّنة)

يُطلق اسم "**العيِّنة**" على أيّة مجموعة من العناصر تُسحب من مجتمع إحصائي بُغية إخضاعها لدراسة إحصائية معيَّنة ومُحدَّدة.

### (٢ , ٣ , ١ , ١) ملاحظات

١- يُدعى عدد عناصر العيّنة بـ **حجم العيّنة** Size of Sample، ويُرمز له بحرف لاتيني صغير من قبيل $n$ أو $m$ أو...

٢- نودّ الإشار ة هنا إلى أنَّه لا يوجد مقياس لتحديد حجم العيِّنة المسحوبة سواء أكان كبيراً أم صغيراً، إذ إنَّ ذلك يتوقف على طبيعة المجتمع المدروس، ونوعية الظاهرة التي هي قيْد الدارسة، وكذلك على الإمكانيات المتوفرة لدى الباحث. في الحقيقية ما زال هناك خلاف في تحديد قيمة حجم العيِّنة $n$ الذي من أجله تكون العيِّنة صغيرة الحجم، فهناك مرجعيات تأخذه $n < 30$ والبعض الآخر $n < 28$ وأخرى تعتمد $n < 27$ (ولكن المراجع التي تعتمد هذه الأخيرة قليلة جداً). إنَّ العدد $n$ الذي يحدِّد كِبَر وصغَر حجم العيِّنة قابل للتعديل أحياناً حسب طبيعة المسألة المدروسة والإمكانيات المتوفرة لدى الباحث والكلفة المترتِّبة على دَراستها، ولكن سنتَّفق على أنَّ العيِّنات التي حجمها $n < 32$ هي عيِّنات صغيرة الحجم، و أمّا العيِّنات التي حجمها $32 \le n$ فهي عيِّنات كبيرة وسنوضِّح سبب ذلك لاحقاً عند بناء جداول التوزيع التكرارية.

٣- من يودّ الاطلاع على المزيد من المعلومات حول العينات وأنواعها وطرائق سحبها يجدها لاحقاً في الفصل التاسع (فصل نظرية المعاينة).

الآن، وبعد إخضاع أفراد المجتمع أو العيِّنة للدراسة الإحصائية ستتولَّد لدينا مجموعة من البيانات الإحصائية ندعوها **بيانات** على سبيل الاختصار، ولذلك يُمكننا أن نصيغ هذا المفهوم من خلال الفقرة الآتية.

### (٤ , ١ , ١) تعريف (البيانات الإحصائية Statistical Data)

**البيانات الإحصائية** (وسنستخدم كلمة "بيانات"على سبيل الاختصار لأنّنا لن نتعامل إلاّ مع بيانات إحصائيّة) هي قياسات Measures أو مَلحُوظات (أو مشاهدات) Observations تهدف إلى غرض معيّن في مجتمع إحصائي مُحدّد، ويتمّ تدوينها نتيجة لـ:

- عمليّة إنتاجية مثل كميّات القمح الناتج عن الزراعة في عام أو أعوام متتالية في بلد ما أو ...
- تجربة معمليّة مثل حساب أعداد جسيمات ألفا الصادرة عن مادة مشعّة خلال فترات زمنية قصيرة (تجارب جايجر-ميرسدن Geiger–Marsden experiments) أو شدّة التيار المار في دارة كهربائية، أو الألوان الناتجة عن تحليل الضوء الأبيض، أو...

- ملاحظات لكائنات موجودة في الطبيعة مثل مراقبة موت الجراثيم (بكتيريا) في مزرعة جرثومية نتيجة لعقار دوائي معيّن، أو مراقبة ولادة واختفاء النجوم في مَجرّة من الفضاء الكوني أو ...

### (١ , ١ , ٤ , ١) ملاحظات

١ - تكون مجموعة بيانات **X** متقطّعة (أو منفصلة) Discrete إذا كانت هذه المجموعة منتهيّة أو غير منتهية ولكنّها قابلة للعدّ، وأمّا إذا كانت هذه المجموعة غير منتهية وغير قابلة للعدّ، فعندئذٍ يُقال عن المجموعة **X** إنّها مستمرّة (أو متصلة) Continuous Set (من أجل الاطلاع على مفهوم المجموعة القابلة للعدّ والمجموعة غير القابلة للعدّ انظر الملحق $A$).

٢ - إنَّ عمليّة جمع البيانات لا تتطلب بالضرورة عناصر مُتخصّصة في علم الإحصاء، إذ يكفي لتأدية هذا العمل تأهيل أفراد مُعيّنين (كالمعلمين في حالات المسح السكاني)، أو المخبريين الذين يقومون بتنفيذ التجارب في المعامل أو...

### (٢ , ١ , ٤ , ١) أمثلة

١ - إنَّ أعداد الأشخاص المهاجرين سنوياً من الأرياف إلى المدن في بلد ما خلال السنوات المتعاقبة تُمثّل مجموعة متقطّعة لبيانات تتعلّق بالهجرة الداخلية للبلد (شهدت الفترة (2007-1957) نزوح نحو 800 مليون نسمة من المناطق الريفية إلى المدن- منظمة الفاو **FAO** Org. 2007).

٢ - أعداد المرضى الذين يتمّ معاينتهم يومياً في مستشفى تمثّل مجموعة متقطّعة لبيانات تتعلّق بالقدرة الاستيعابية للمستشفى.

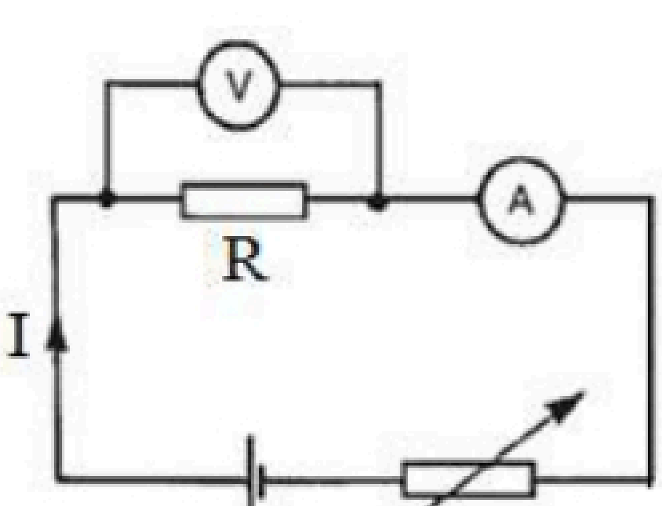
٣- تَغيُّر مستويات الضغط مع مرور الزمن داخل مرجَل بخاري (عال جداً- عال- عادي- منخفض- منخفض جداً) تمثّل مجموعة متقطّعة لبيانات تُعبّر عن تَغيّر الضغط داخل هذا المرجَل.

٤ - لنأخذ تجربة إلقاء حجر نرد Die (سداسي الوجوه) وقذف قطعة نقود معدنية Coin لمرّةواحدة فقط.

فعندئذٍ سيكون لدينا مجموعة النتائج الممكنة الآتية:

| | | | | | |
|---|---|---|---|---|---|
| $(H,1)$ | $(H,2)$ | $(H,3)$ | $(H,4)$ | $(H,5)$ | $(H,6)$ |
| $(T,1)$ | $(T,2)$ | $(T,3)$ | $(T,4)$ | $(T,5)$ | $(T,6)$ |

وهي تمثّل مجموعة منتهية من البيانات المتعلّقة بهذه التجربة، علماً أنَّ الحرف $T$ يشير إلى الشعار (أول حرف من **Tails**) و$H$ يشير إلى الصورة (أول حرف من **Heads**)، ومن ثمَّ نلاحظ أنَّ مجموعة البيانات التي حصلنا عليها هي مجموعة متقطّعة.




٥- لنأخذ تجربة تغيُّر شدّة التيار الكهربائي **I** المار في مقاومة أوميَّة Resistance رمزها **R** لدى تغيير فرق الجهد V على طرفيها، فإنَّ القيم التي سنحصل عليها هي مجموعة مستمرَّة لبيانات تتعلّق بهذه التجربة.

٦- لدى تتبع حالة الطقس في منطقة ما فإنَّ القيم العدديّة النّاتجة عن قياسات أجهزة الرصد كالحرارة والرطوبة والضغط وسرعة الرياح هي بيانات تنتمي إلى مجموعات مستمرَّة لبيانات تتعلّق بحالة الطقس.

٧- لو أخذنا الأرقام الجامعية (ID) لمجموعة من الطلاب (حيث لكل طالب جامعي رقم خاص يميّزه عن بقية الطلاب)، فإنَّ هذه الأرقام تكوِّن مجموعة متقطِّعة لبيانات تتعلَّق بهؤلاء الطلاب.

إذن، وكما نلاحظ، فإنَّ جميع ما سبق هي حقائق أو مشاهدات أو نتائج لتجارب كما أسلفنا في تعريف البيانات، ولكن السؤال الذي يطرح هنا هو: **كيف يمكن لهذه البيانات أن تَنْتج لدينا، بمعنى ما هي الوسائط التي تُولِّد لنا هذه البيانات؟**

إنَّ الإجابة على هذا السؤال تكمن في الفقرة التالية.

## (١,٢) المتغيرات وأنواع البيانات

**Variables and Classification of Data**

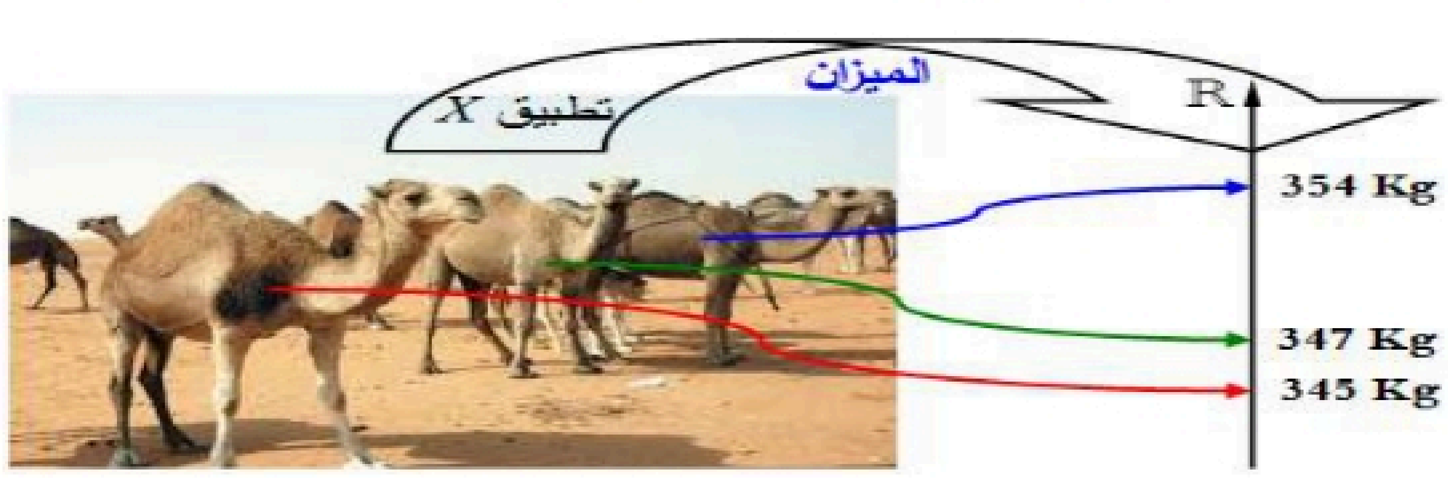
نقدِّم فيما يلي أحد أهمّ المفاهيم المتعلِّقة بتوليد البيانات.

### (١ , ٢ , ١) المتغيِّرات Variables

في الواقع تنتج البيانات (سواء كانت ناتجة عن قياسات أو استقصاءات ) عن وسيلة تُطَّق على عناصر العيِّنة (أو المجتمع في حال المسح الشامل) تُدعى مُتغيِّراً Variable، والأمثلة الآتية توضِّح لنا جانباً من ذلك.

#### (١ , ١ , ٢ , ١) أمثلة

١- لنفترض أنَّنا بصدد تنفيذ دراسة لمعرفة متوسط الوزن لمجموعة من الإبل في مزرعة ما، فعندئذ نأتي بميزان مناسب لهذه الغاية، ومن ثمَّ نخضع كل رأس من الإبل يقع عليه الاختيار للوزن فنحصل نتيجة لذلك على مجموعة من الأرقام ولتكن 354 و347 و345 و...، وفي هذه الحالة نلاحظ أنَّ الميزان الذي ولَّد البيانات قد لعب دور المتغيِّر. لذلك يُقال في مثل هذه الدراسات إنَّ ظاهرة الوزن (أو الوزن على سبيل الاختصار ) هي متغيِّر لأنَّه كلما طبِّق على رأس من الإبل أعطى قيمة (حقيقية). لاحظ هنا أنَّ هذا المتغيِّر هو تطبيق Map مجاله (أو مجموعة تعريفه أو نطاقه) Domain مجموعة الإبل في حين أنَّ مجاله المقابل (أو مستقره أو نطاقه المرافق) Co-domain هي فترة من مجموعة الأعداد الحقيقية (من أجل الاطلاع على هذا المفهوم انظر الملحق A).




٢- لنفترض أنَّنا بصدد تنفيذ دراسة لمعرفة أعداد أفراد الأسر لطلاب جامعة الملك سعود، فعندئذ إمّا أن يُقدَّم استبيان أو يُسأل الطّالب مباشرةً أو يُراسل بطريقة ما ليجيب على السؤال المتعلِّق بهذه الدراسة: **كم عدد أفراد أسرتك؟** فنحصل نتيجة لذلك على مجموعة من الأرقام (بيانات) ولتكن 7، 12، 6 و...، وفي هذه الحالة نلاحظ أنَّ السؤال السابق لعب دور المتغيِّر لأنَّه كلما طبّق على طالب أعطى قيمة (بيان)، وكما هو ملاحظ فإنَّ هذا المتغيِّر هو تطبيقٌ مجاله مجموعة الطّلاب، وأمّا مجاله المقابل فهو مجموعة جزئية من مجموعة الأعداد الطبيعية.

٣- لنفترض أنَّنا بصدد تنفيذ دراسة لمعرفة ألوان العيون في مجتمع بشري ما، فعندئذ نأتي بجهاز لتحديد درجة اللون (أو خبير بتمييز ألوان العيون)، ومن ثمَّ نخضع كل شخص يقع عليه الاختيار للقياس بهذا الجهاز فنحصل نتيجة لذلك على مجموعة من الصفات (بيانات) بني، أسود، أخضر و...، وفي هذه الحالة نلاحظ أنَّ جهاز تحديد اللّون الذي يولِّد البيانات الممثِّلة لظاهرة لون العيون لعب

دور المتغيِّر. لذلك يكون الجهاز الذي حدَّد لون العيون هو المتغيِّر، وذلك لأنَّه كلما طبّق على شخص أعطى بياناً ( صفة أو مُسمّى)، وهو تطبيقٌ مجاله مجموعة الأشخاص ومجاله المقابل هي مجموعة الألوان.

٤ – لنفترض أنّنا بصدد تنفيذ دراسة تتعلّق بتقدير اجتهاد الطّلاب في مقرّر دراسي لدى إحدى الكليات في جامعة الملك سعود (بطريقة المسح الشّامل)، فعندئذ نُخضع كل طالب في هذه الكلية للاختبار، وبعد عملية التصحيح والرصد نحصل على مجموعة من التقديرات (بيانات) مثل $A^+$ وَA و$B^+$ وB و$C^+$ وC و$D^+$ وD وF، وفي هذه الحالة نلاحظ أنَّ نموذج الأسئلة الذي قدِّم للطلاب لعب دور المتغيِّر. لذلك يُقال في مثل هذه الدراسات إنَّ نموذج الأسئلة هو المتغيِّر لأنَّه كلما طبق على طالب أعطى بياناً (رمزاً)، وهو تطبيقٌ مجاله مجموعة الطلاب ومجاله المقابل مجموعة التقديرات.

أمّا عن سبب تسميتها متغيِّرات فإنَّ ذلك يعود إلى نتائج تطبيقها على العناصر، فكلما طُبّق على عنصر أعطى نتيجة قد تختلف من عنصر إلى آخر، إذ إنَّ تلك القيم ليست بالضرورة أن تكون متساوية أو متماثلة، وإن كانت متساوية أو متماثلة فلا ضير من ذلك، وبناءً على ما سبق نلاحظ أنّه من منطلق رياضيّاتي أنَّ المتغيِّر هو تطبيق (يمثِّل وسيلة أو سمة) مجاله مجموعة عناصر العيِّنة (أو المجتمع في حال المسح الشامل) ويأخذ قيمه في مجموعة ذات طبيعة ما، ولذلك يُمكننا أن نصيغ مفهومَ المتغيِّر كما في التعريف الآتي.

**(٢ , ١ , ٢ , ١) تعريف (المتغيِّر)**

المتغيِّر هو تطبيق يمثِّل وسيلة أو سمة (أو أي مولِّد للبيانات) مجاله العيِّنة (أو المجتمع في حال المسح الشامل) ويأخذ قيمة في مجموعة ذات طبيعة ما (قد تكون عددية أو غير ذلك).

إذا كانت مجموعة قيم المتغيِّر قابلةً للعدِّ على الأكثر فعندئذ يُقال عنه إنَّه متغيِّر متقطِّع Discrete Variable، وأمّا إذا كانت مجموعة قيم المتغيِّر غير قابلةً للعدِّ فعندئذ يُقال عنه إنَّه متغيِّر مستمرّ Continuous Variable.

لقد لاحظنا أنَّ تطبيق المتغيِّرات على عناصر العيِّنة أو المجتمع يولِّد لنا البيانات، وهذه البيانات ليست من ذات الطبيعة (أو من النوع نفسه)، فلو أمعنا النظر في طبيعة البيانات التي حصلنا عليها في الأمثلة السّابقة نجد أنَّها ليست من طبيعة واحدة، فمنها ما هو رقمي ناتج عن قياس، والآخر وصفي (رمزي أو اسمي، وقد يكون رقمياً ولكنَّه غير ناتج عن قياس) وقد تكون مزيجاً. في الواقع يمكن أن تكون البيانات المصنَّفة إمّا كميّة أو نوعيّة على أبعد حد، وذلك لأنَّ البيانات التي من النوع المزيج تندرج تحت اسم البيانات النوعيّة أيضاً. لذلك سنقوم فيما يلي بتقديم تصنيف البيانات بُغية تسهيل دراستها.

**(٢ , ٢ , ١) تعريف ( البيانات النوعيّة Qualitative Data)**

البيانات النوعيّة (أو البيانات الوصفيّة) هي تلك البيانات التي تنتج عن التساؤل بـ "ما" أو "كيف".

تتكون البيانات النوعيّة عادةً من:

١ – أسماء مثل: صورة وشعار في رمي قطعة النقود.

٢ – صفات مثل: بني، أسود، أزرق، وأخضر في دراسة لون العيون.

٣ – أرقام لها دور الصفات مثل: الأرقام الجامعية (ID) للطّلاب في جامعة ما.

بكلمات أخرى. إنَّ البيانات النوعيّة تتكوَّن بشكل أساسي من تسميات Labels أو أسماء Names أو أرقام مستخدمة لتحديد سمة كل عنصر من العيِّنة أو المجتمع، وغالباً ما يُشار إليها بـ البيانات الفئوية، كأن يُقال مثلاً: فئة الأشخاص ذوي العيون السود، أو فئة ذوي الدخل المحدود، فلو رجعنا إلى الأمثلة (٢ , ٤ , ١ , ١) السّابقة فإنَّنا نجد بيانات الفقرات / ٣، ٤ و٧/ هي بيانات نوعيّة، وكذلك الفقرتين / ٣ و٤/ من الأمثلة (١ , ١ , ٢ , ١) هي بيانات نوعيّة أيضاً، والمزيد من الأمثلة حول البيانات النوعيّة تُقدمها لنا الفقرة الآتية.

**(١ , ٢ , ٢ , ١) أمثلة**

١- الجنسيّة لمجموعة من الأشخاص المقيمين في بلد ما يولّد بيانات من قبيل: آسيوي، أوربي، أمريكي، أسترالي و...، وهي بيا نات نوعيّة لأنّها مُسميّات لها دور الصفات. لاحظ هنا أنَّ هذا النوع من البيانات لا يقبل صفة الترتيب، والترتيب الذي وردت به لا يعني أكثر من وجود لمثل هذه الجنسيات فقط.

٢- الحالة الصحية لمرضى في مستشفى تولّد بيانات من قبيل: ممتازة (متماثل للشفاء)، جيدة جداً (في تحسن مستمر)، جيدة (وضعه مستقر ولا توجد لديه مضاعفات)، مقبولة (في وضع غير مستقر)، سيئة (وضعه غير مستقر ولديه مضاعفات)، وهي تقديرات لها صفة البيانات النوعيّة. لاحظ هنا أنَّ هذا النوع من البيانات يقبل صفة الترتيب فيما بينها، والترتيب الذي وردت به له معنى.

٣- الحالة الاجتماعية لأفراد مجتمع بشري يولّد بيانات من قبيل: عازب، متزوِّج، مطلَّق وأرمل، وهي مسميّات أو صفات، ومن ثَّم فهي جميعاً بيانات نوعيّة. لاحظ هنا أنَّ هذا النوع من البيانات لا يقبل صفة الترتيب، والترتيب الذي وردت به لا يعني أكثر من وجود لمثل هذه الحالات في المجتمع البشري.

٤- حالة الطقس في منطقة ما من الأرض يولّد بيانات من قبيل: صحو، غائم، ماطر، مثلج، مُغبَّر و...، وهي مسميّات أو صفات، ومن ثَّم فهي جميعاً بيانات نوعيّة. لاحظ هنا أنَّ هذا النوع من البيانات لا يقبل صفة الترتيب أيضاً، والترتيب التي وردت به لا يعني أكثر من وجود مثل هذه الحالات في المُناخ فقط.

**(٢ , ٢ , ٢ , ١) ملاحظات**

١- إنَّ المتغيَّرات التي تولّد بيانات نوعيَّة تُدعى **متغيِّرات نوعيَّة** Qualitative Variables، ومن ثَّم تكون المتغيِّرات التي وردت في الفقرات / ٣، ٤ و٧/ من الأمثلة (٢ , ٤ , ١ , ١) وكذلك الفقرتين / ٣ و٤/ من الأمثلة (١ , ١ , ٢ , ١) وجميع فقرات المثال (١ , ٢ , ٢ , ١)، هي متغيِّرات نوعيَّة متقطِّعة (باستثناء المتغيِّر المولِّد لبيانات المثال / ٣ / من (١ , ١ , ٢ , ١) يمكن له أن يكون مستمرّاً)، ونشير هنا إلى أنَّ مجموعة قيم هذه المتغيِّرات قد تكون غير قابلة للعدّ (كما في طيف الضوء المرئي).

٢- يمكن تصنيف البيانات النوعيّة وفقاً لمقاييس رئيسيّة عديدة هي:

**أ- المقياس الاسمي Nominal Scale**

إنَّ البيانات النوعيّة التي تنطوي تحت هذا النوع من المقاييس تكون غير قابلة للترتيب فيما بينها، وكذلك تكون التصنيفات في هذا النوع من القياسات مختلفة وغير متكرِّرة وليس لها أيّة دلالات عدديّة تقبل الحساب، وإن كانت البيانات من النوع الرقمي فهي لتسهيل التعامل مع التصنيفات ولا تُعطيها أيّة إمكانية للترتيب، فعلى سبيل المثال:

أ-١- لو أخذنا البيانات النّاتجة عن قذف قطعة نقود معدنيّة لمَّرة واحدة فقط حيث لدينا $H$ (صورة) و$T$ (شعار)، فنلاحظ أنَّه ليس لدينا ما يفضّل إحداهما على الأخرى في الترتيب.

أ-٢- لو أخذنا البيانات النّاتجة عن دراسة لون العيون عند البشر حيث لدينا بنّي، أسود، أزرق و...، فنلاحظ هنا أنَّه ليس لدينا ما يفضّل لون على آخر في الترتيب.

أ-٣- لو أخذنا المهام الموكلة إلى أعضاء فريق كرة قدم حيث لدينا المهاجم (ولنفترض أنَّ رقمه 10)، الجناح الأيمن (ولنفترض أنَّ رقمه 17)، الجناح الأيسر (ولنفترض أنَّ رقمه 15)، الدفاع (ولنفترض أنَّ رقمه 22) وحارس المرمى (ولنفترض أنَّ رقمه 1 )، فعندئذ لا يُمكننا القول إنَّ اللاعب رقم 10 أكثر مهارة من اللاعب رقم 1، أو أنَّ اللاعب رقم 22 أكثر مهارة من اللاعب رقم 15، فلكلٍّ لاعب من هؤلاء مهارات وفنون لعب ومهام تختلف عن الآخر، ومن ثَّم عملية الترتيب بين الأرقام هنا ليست ذات معنى.

أ-٤- لو أخذنا البيانات النّاتجة عن تحديد الحالة الاجتماعية (عازب، متزوج، مطلق، أرمل) لمجتمع حيّ من أحياء مدينة ما، أو

البيانات النّاتجة عن تحديد حالة الجنس لمجموعة نباتات في غابة (مذكر، مؤنث، خنثى )، فإنَّنا نجدها بيانات نوعيّة تفتقر إلى خاصيّة الترتيب فيما بينها أيضاً.

**ب- المقياس الترتيبي Ordinal Scale**

إنَّ البيانات النوعيّة التي تنطوي تحت هذا النوع من المقاييس تكون قابلة للترتيب فيما بينها وفقاً لمعيار مُحدَّد وواضح، كأن تكون من الأقوى إلى الأضعف أو من الأعلى إلى الأدنى أو... ، فعلى سبيل المثال:

**ب-١**- لنفترض أنَّنا بصدد دراسة قساوة المعادن، وأخذنا مجموعة من المعادن مكوَّنة من الحديد، والنحاس، والألماس، والذهب، والرصاص، وأردنا أن نعيِّن أيها أكثر قسوةً من الآخر (**باستخدام الخدش فيما بينها، وقد وضع هذا المقياس عام 1812 م على يد عالم المعادن الألماني موس Mohs** )، فنجد أنَّ الألماس يأتي في الترتيب الأول ويليه الحديد، فالنحاس، ومن ثَّم الذهب وأخيراً الرصاص، فنلاحظ أنَّ هذه البيانات النوعيّة (**الألماس، والحديد، والنحاس، والذهب، والرصاص**) أمكن ترتيبها وفقاً لمعيار القساوة.

**ب-٢**- لو أخذنا نظام تقدير نتائج الاختبارات في الجامعات السعودية وهي $A^+$، $B^+$،$C^+$،$D^+$،A، B، C، D و F فنجد أنَّ هذه البيانات النوعيّة لها ترتيب تنازلي $A^+$، A، $B^+$، B، $C^+$، C، $D^+$، D و F التي تعني أنَّ الطالب الذي حصل على تقدير B (**أي إنَّه حصل على درجة $x$ مع** $80 \le x \le 84$) أقل من التقدير السابق له $B^+$ (**الحاصل على درجة $x$ مع** $85 \le x \le 89$).

**ب-٣**- لو أخذنا استبياناً مقدَّماً للطلاب من أجل دراسة إحصائية تخص المحاضرين حيث يُطلب من الطالب الإجابة على مجموعة من الأسئلة بإحدى العبارات المدوَّنة في الجدول الآتي:

| المعيار | أوافق بشدة | أوافق | محايد | لا أوافق | لا أوافق مطلقاً |
|---|---|---|---|---|---|
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |

فنلاحظ أنَّ البيانات الناتجة عن هذا الاستبيان (أوافق بشدّة، أوافق، محايد، لا أوافق ولا أوافق مُطلقاً) ، هي بيانات نوعيّة تقبل الترتيب إلاَّ أنَّ المشكلة في تفسير نتائج تحليلها تكمن في عدم وضوح التباعد بين هذه البيانات، فمن الممكن أن يجيب طالب بعبارة "لا أوافق" وآخر يستخدم عبارة "لا أوافق مطلقاً" على معيّار ما، وعندما نستفسر منهما كلٌ على انفراد ما يعنيه بإجابته نجد لديهما التصوّر نفسه حول الإجابة على هذا المعيّار، بمعنى أنَّ كلاً منهما استخدم بُعداً خاصاً به لفهم عبارة "**لا أوافق**" و "**لا أوافق مطلقاً**". لذلك يُفضّل في مثل هذه الحالات استخدام تقييمات رقمية لتحديد التباعد بين كلّ تقييم وآخر.

**ب-٤**- بالرجوع إلى الفقرة / ٣/ من الأمثلة (٢ , ٤ , ١ , ١) فإنَّنا نجد أنَّ البيانات الناتجة عن تحديد مستوى الضغط في المرجل هي بيانات ترتيبيّة، وإذا كان معيّار الترتيب من الأعلى ضغطاً إلى الأقل فيكون لدينا الترتيب الآتي لتلك البيانات: عال جـداً– عال– عادي– منخفض– منخفض جداً، وهكذا نجد أنَّ هذه البيانات تتبع المقياس الترتيبي.

**ب-٥**-لو أخذنا الأرقام الجامعيّة لطلاب إحدى كليّات جامعة ما، فعندئذ نلاحظ أنَّ هذه البيانات النوعيّة يمكن ترتيبها وفقاً لتسلسل قيم الأعداد من الأكبر إلى الأصغر أو بالعكس، ولكن قد يكون هذا الترتّيب غير ذي معنى إذا بدأت من الصفر ولفصل مُحدَّد (وتكون البيانات في هذه الحالة اسميّة). لكن يمكن أن تصبح ذات معنى إذا ما حوت الأرقام على ما يدل على عام التسجيل للطالب مثلاً كما هو الحال لدى طلاب جامعة الملك سعود حيث خصّصت الخانات الثلاث اليسرى من الرقم الجامعي لتشير إلى العام الذي سُجّل فيه الطالب في الجامعة، فعلى سبيل المثال نجد أرقاماً من النماذج ...433،...434، ...435،...436 و...، وفي هذه الحالة يمكننا ترتيب البيانات في مجموعات وفقاً لعام التسجيل، ولكنّها ضمن المجموعة السنويّة الواحدة هي بيانات اسميّة وليست ترتيبية.

**(٣ , ٢ , ١) تعريف (البيانات الكمّية Quantitative Data)**

البيانات الكمّية هي بيانات تتكون من أرقام تمثّل قيماً عددية تنتج عن التساؤل بـ " كم".

غالباً ما تكون البيانات الكمّية ناتجة عن قياسات أو إحصاءات أو تعدادات، ومن كونها قيماً عدديّة فإنَّ العمليّات الحسابية العاديّة ذات مغزى عند التعامل مع هذا النوع من البيانات.

لاحظ أنَّ الفقرات / ١، ٢، ٥ و٦/ من الأمثلة (١-١-٤-٢) وكذلك الفقرتين / ١ و٢/ من الأمثلة (١-٢-١-١) تُقدِّم لنا بيانات كمّية.

**(١ , ٣ , ٢ , ١) ملاحظات**

١- إنَّ المتغيِّرات التي تولِّد بيانات كميَّة تُدعى **متغيِّرات كمِّية** Quantitative Variables، ومن ثَّم تكون المتغيِّرات الواردة في الفقرات / ١ و٢ و٥ و٦/ من الأمثلة (٢ , ٤ , ١ , ١) وكذلك الفقرتين / ١ و٢/ من الأمثلة (١ , ١ , ٢ , ١) هي متغيِّرات كمِّية.

٢- بناءً على ما سبق بخصوص قدرة مجموعة قيم متغيِّر، فإنَّ المتغيِّرات الواردة في الفقرتين/ ٥ و٦/ من الأمثلة (٢ , ٤ , ١ , ١) وكذلك الفقرة / ١/ من الأمثلة (١ , ١ , ٢ , ١) هي متغيِّرات مستمرّة، في حين أنَّ المتغيِّرات الواردة في الفقرتين / ١و٢/ من الأمثلة (٢ , ٤ , ١ , ١) وكذلك الفقرة / ٢/ من الأمثلة (١ , ١ , ٢ , ١) هي متغيِّرات متقطِّعة.

٣- يمكن تصنيف البيانات الكمّية في مقياسين رئيسين أيضاً هما:

**أ- مقياس الفترة Interval Scale**

هذا النوع من المقاييس استمد اسمه من كون البيانات التي تتبعه تنتمي لفترة من مجموعة الأعداد الحقيقية، وله خصائص أخرى كالمحافظة على خاصّية الترتيب للبيانات، بمعنى أنّه يُمكن ترتيبها؛ وذلك لأنَّ العلاقة ($\le$) هي علاقة ترتيب جزئي على مجموعة الأعداد الحقيقية $\mathbb{R}$ ولكن من أجل هذا النوع من البيانات لا معنى لعمليّة نسْبة الكميات بعضها إلى بعض، وكذلك لا تعتمد قيمة الصفر كنقطة بداية له إن وجدت (بمعنى أنَّ الصفر هنا لا يعني العدم). إذاً يُمكننا أن نميِّز مقياس الفترة بالخصائص الآتية:

**أ-١** - قيمته تنتمي إلى فترة من $\mathbb{R}$ (قيم البيانات الناتجة عن هذا المقياس تنتمي لمجموعة غير قابلة للعدّ).

**أ-٢** - يمكن ترتيب قيم البيان.

**أ-٣**- يمكن قياس درجة الاختلاف بين البيانات، ولكن دون إمكانية استخدام النسبة فيما بينها، بمعنى أنَّ عملية نسبة الكميات بعضها إلى بعض غير ذي معنى أو ليس له دلالة واضحة.

**أ-٤** - قيمة الصفر في هذا النوع من البيانات لا تعني العدم أو عدم امتلاك أي شيء (انظر الأمثلة الآتية)، بمعنى أنَّ هذا النوع من البيانات لا يعتمد الصفر كنقطة بداية (أو نقطة أصل).

**ب- مقياس النسبة Ration Scale:**

لقد سمِّي هذا المقياس بهذا الاسم لأنَّه تقدير نسبة بين كمّية وأخرى من النوع نفسه، وعلاوة على ذلك فإنَّ هذا المقياس يحافظ على جميع الخصائص التي يتمتَّع بها مقياس الفترة، ويزيد على ذلك أنّه من الممكن التعامل مع بيانات لها قيم متقطِّعة أيضاً، وأخيراً من مميِّزاته الفارقة عن المقاييس الأخرى أنَّ قيمة الصفر هنا إن وجدت فلها معنى واضح وتُعتمد كنقطة بداية (أو نقطة أصل).

إذاً يمكننا تمييز مقياس النسبة بامتلاكه لخصائص مقياس الفترة مضافاً إليها الخصائص الآتيّة:

ب-١- قيم البيانات قد تنتمي إلى فترة (مجموعة مستمرَّة) أو إلى مجموعة قابلة للعدّ على الأكثر.

ب-٢- من الممكن نسبة الكميات بعضها إلى بعض، وعمليّة النسبة هذه لها معنى.

ب-٣- يُعتمد الصفر كنقطة بداية، وهذا يعني أنَّ قيمة الصفر إن وجدت لها معنى واضح (قيمة الصفر قد تعني العدم).

**(٢ , ٣ , ٢ , ١) أمثلة**

١- المتغيّر الذي يرصد الزمن اللازم لإنهاء كتابة واجب منزلي من قبل طالب (مقدراً بالدقيقة) هو متغيِّر كميّ مستمرّ تبدأ قيمه بعددموجب تماماً وتمتد إلى فترة محدودة فتصبح مجموعة قيمه من الشكل $]0,\alpha)$، والصفر هنا غير موجود؛ لأنَّه يعني أنَّ الطالب لم يبدأ بكتابة الواجب أصلاً. من جهة أخرى يُمكننا هنا نسبة الكميات بعضها إلى بعض كأن نقول: إنَّ الزمن اللازم لكي يُنهي طالب كتابة واجبه المنزلي لمقرَّر الرّياضيّات يساوي ضعف الزمن اللازم لكي ينهي ذلك الطالب كتابة واجبه المنزلي لمقرَّر الإحصاء، ومن ثَّم البيانات التي نحصل عليها من هذا المتغيّر تتبع مقياس النسبة.

٢- المتغيّر الذي يرصد عدد العاطلين عن العمل في مدينة ما خلال أشهر متعاقبة هو متغيِّر كميّ متقطِّع قد تبدأ قيمه من الصفر وتمتد إلى أعداد صحيحة غير سالبة ولكنَّها محدودة، فتصبح مجموعة قيمه من الشكل 0 و1 و2 و... و$n$ علماً أنَّ $n$ عدد صحيح محدود. من جهة أخرى يمكننا هنا نسبة الكميات بعضها إلى بعض، كأن نقول إنَّ عدد العاطلين عن العمل في شهر شباط (فبراير) من هذا العام يساوي نصف عدد العاطلين عن العمل في شهر شباط من هذا العام الماضي. أخيراً نجد أنَّ للصفر هنا معنى واضح، فإذا كان عدد العاطلين عن العمل في شهر شباط من هذا العام يساوي 0 فهذا يعني أنَّه لم يَعد لدينا عاطلون عن العمل في هذا الشهر، ومن ثَّم البيانات التي نحصل عليها من هذا المتغيِّر تتبع مقياس النسبة أيضاً.

٣- المتغيّر الذي يرصد درجة حرارة الطقس في منطقة معيَّنة من الأرض خلال الأيام المتعاقبة من عام معيَّن مُقدرة بالدرجة المئوية Celsius ويرمز لها بـ $^{o}\mathbf{C}$، أو الفهرنهايت Fahrenheit ويرمز لها بـ$^{o}\mathbf{F}$ علماً أنَّ $^{o}\mathbf{F} := 1.8(^{o}\mathbf{C}) + 32$ هو متغيِّر كميّ مستمرّ، والقيم التي نحصل عليها قد تمتد من القيمة السالبة 60- درجة مئوية (كما يحصل في شرق كازاخستان حيث تمَّ تسجيل (60-) درجة مئوية) وحتى 71+ درجة مئوية (حيث سجل أحد الأقمار الاصطناعية (71+) درجة مئوية في صحراء لوط بإيران)، وهنا نلاحظ أنَّ الصفر لا يعني أنَّه لا توجد حرارة وإنَّما الطقس بارد إلى حد درجة تجمد الماء، ومن ثَّم لا ينظر إلى درجة الصفر على أنَّها نقطة بدء وإنَّما درجة يمكن بلوغها. من جهة أخرى لا يمكننا نسبة هذه المقادير بعضها إلى بعض، فمن غير الجائز أن نقول إنَّ درجة حرارة منطقة **A** تساوي ضعف درجة حرارة منطقة أخرى **B**، وإنَّما يمكننا القول إنَّ الفرق بين درجتي الحرارة للمنطقتين يساوي $x$ مثلاً. فعلى سبيل المثال إذا كانت درجة الحرارة في المنطقة **A** تساوي 0 درجة مئوية (وهي تعني أنَّ هذه المدينة باردة نسبياً) وكانت درجة الحرارة في المنطقة **B** تساوي 20 درجة مئوية (وهي تعني أنَّ هذه المدينة دافئة نسبياً)، فعندئذ لا نستطيع القول إنَّ المنطقة **B** أدفأ عشرين مرَّة من المدينة **A** ولا القول إنَّ المنطقة **A** أبرد عشرين مرَّة من المنطقة **B**، وكل ما يمكننا قوله هو إنَّ الفرق بين درجتي الحرارة للمنطقتين يساوي 20 درجة مئوية. إذاً درجة الحرارة في هذا المثال هي من النوع الفتراتي، ومن ثَّم البيانات التي نحصل عليها من هذا المتغيِّر تتبع مقياس الفترة.

٤- المتغيّر الذي يرصد درجة الحرارة لغاز مثالي مُقدرةً بـ الكلفن Kelvin (ويرمز لها بـ **K**)، علماً أنَّ $\mathbf{K} := {}^{o}\mathbf{C} + 273.15$ هو متغيِّر كميّ مستمرّ، وقيمته قد تمتد من 0 كلفن (تُدعى الدرجة $\mathbf{K}=0$ بالصفر المطلق، وهي ما تزال في إطار درجات الحرارة النظرية التي لم يتمّ الوصول إليها بعد) إلى أي قيمة موجبة أخرى ولكنَّها محدودة. فإذا ما وصلت درجة حرارة الغاز المثالي إلى الصفر كلفن فحينئذ ينعدم الضغط لهذا الغاز، والصفر هنا هو نقطة بدء للقياس، ومن ثَّم البيانات التي نحصل عليها من هذا المتغيِّر تتبع مقياس النسبة.

٥- المتغيّر الذي يرصد المسافة بين مدينة **A** وأخرى **B** هو متغيِّر مستمرّ، ويأخذ قيمه في فترة $[0,\alpha)$ مع $\alpha$ عدد حقيقي محدود. والصفر هنا غير موجود بسبب التمايز بين المدينتين فرضاً، وفي حال وجوده يكون له معنى، فإذا كانت المسافة تساوي الصفر فإنَّ ذلك يعني أن المدينة **A** هي **B** نفسها. كذلك نلاحظ أنَّه يمكننا تنفيذ عملية النسبة على المسافات بين المدن، كأن نقول إنَّ البعد بين المدينة **A** و**B** يساوي ضعف المسافة بين **A** و**C**، وبذلك تصبح البيانات التي نحصل عليها من هذا المتغيِّر تتبع مقياس النسبة.

٦- المتغيّر الذي يرصد أوزان صناديق تحوي أناناس، ومفترضين أنَّ وزن الصندوق الذي سيحتوي هذه الفاكهة مهملاً بالمقارنة مع وزن حبة الأناناس (بمعنى أنَّ دقَّة الميزان لن تظهر وزناً للصندوق الفارغ)، فنجد أنَّ مقياس أوزان هذه الصناديق هو متغيِّر مستمرّ ويأخذ قيمه في فترة $[0,\alpha)$ مع $\alpha$ عدد حقيقي محدود، والصفر هنا يعني أنَّ الصندوق وجد فارغاً. كذلك نلاحظ أنَّه يمكننا تنفيذ عملية النسبة على هذه الصناديق حيث يُمكننا القول إنَّ وزن الصندوق رقم 3 يساوي مرَّة ونصف من وزن الصندوق ذي الرقم 5 مثلاً، وهكذا نلاحظ أنَّ البيانات الناتجة عن هذه الأوزان تتبع مقياس النسبة. لكن يجب الانتباه عند التعامل مع متغيِّر الوزن، وذلك لأنَّ الصفر مرتبط بدقة الجهاز الذي يقيس الوزن حيث نعلم أنّه لا يوجد شيء على الأرض ليس له وزن بسبب جاذبية الأرض، فعلى سبيل المثال لو قمنا بوزن الصندوق باستخدام ميزان دقته غرام واحد فإنَّنا على الأرجح لن نحصل على الصفر حتى لو كان الصندوق فارغاً، في حين أنَّنا لو قمنا بوزنه في ميزان آخر دقته 1000 كغ (طن) فإنَّ الميزان سوف يشير إلى قيمة الصفر حتى وإن كان الصندوق مليئا بالأناناس. بالطبع هذه حالات للمناقشة فقط، وأمّا في الجانب التطبيقي فإنَّه يستخدم من أجل كل دراسة مقياساً أو معياراً مناسباً للدراسة الإحصائيّة المطلوبة.

### (١,٢,٣,٣) ملاحظات

١- نشير هنا إلى أنَّ الطول، العرض، الارتفاع والعمق هي حالات خاصّة من مفهوم المسافة، ويعاملان معاملة المسافة التي سبق وصفها.

٢- يمكن رد البيانات النوعيّة إلى بيانات كميّة من خلال إعطاء قيمة عددية تميّز الاسم أو الصفة التي ولّدت البيان النوعي الناتج عن التجربة أو الدراسة، كأن نعطي في الدراسة المتعلّقة بمستوى الضغط في مرجل بخاري القيم 5، 4، 3، 2 و1 عوضاً عن مرتفع جداً، مرتفع، عادي، منخفض، منخفض جداً على الترتيب، أو أن نعطي في الدراسة المتعلّقة بقذف قطعة نقود معدنية القيم 0 و1 عوضاً عن $H$ و$T$، ولكن على ألاَّ نخضع هذه البياناتالأخيرة للعمليات الحسابية التي تطبَّق على البيانات الكميّة. لذلك ستكون دراساتنا القادمة مركّزةً على البيانات الكميّة، وهذا التخصيص لا يقلّل من شأن وعموميّة هذه الدراسة.

## (١,٣) تمثيل البيانات وعرضها

**Representation and Graphing of Data:**

بعد إتمام جمع البيانات تأتي مرحلة البحث في كيفية التعامل مع هذه البيانات، إذ إنّه يتمّ التحقُّق أولاً من اكتمال عدد البيانات، فإذا كانت هناك بعض البيانات المفقودة فحينئذ يجب النظر في كيفية تحصيلها ثانية أو تقديرها، علماً أنَّ عمليّة التقدير للبيانات المفقودة تقع خارج إطار هذا الكتاب ولذلك لن نتطرّق إلى عملية تقدير البيانات المفقودة. إنَّ البيانات التي نحصل عليها تكون عادة على أشكال مختلفة، فمنها على شكل قيم عددية مفردة، وبعضها الآخر يعرض تَغيُّر ظاهرة ما مع مرور الزمن أو مع مسمَّيات كالبلدان، أو المدن، أو مع كليهما معاً، ومجموعات البيانات الآتيّة توضّح لنا بعضاً من هذه النماذج، حيث سنستشهد بها بين الحين والآخر.

**مجموعة البيانات (٣, ١ .أ ) (بيانات عيِّنة على شكل قيم عددية مفردة):**

تمثّل البيانات الآتية أعمار 64 بطارية جهاز حاسب محمول من نوع مُحدّد (مقدّرةً بالشهر):

| | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 11 | 4 | 13 | 15 | 12 | 4 | 13 | 12 | 17 | 5 | 13 | 11 | 15 | 12 | 5 | 12 | 5 | 13 |
| 6 | 12 | 11 | 6 | 12 | 6 | 6 | 12 | 7 | 15 | 12 | 7 | 13 | 11 | 13 | 7 | 7 | 12 | 7 |
| 9 | 11 | 12 | 9 | 11 | 9 | 15 | 12 | 9 | 11 | 9 | 12 | 9 | 11 | 9 | 15 | 17 | 19 | 17 |
| 11 | 11 | 13 | 13 | 15 | 13 | 15 | | | | | | | | | | | | |

**مجموعة البيانات (٣, ١ .ب ) (بيانات عيِّنة على شكل قيم عددية مرتبطة بالزمن):**

لتكن لدينا البيانات الآتية لعدد الحجيج من خارج المملكة العربية السعودية منذ عام 1423 وحتى عام 1432 هـ (مأخوذة عن موقع الإحصاءات العامة في المملكة العربية السعودية):

الجدول (١ , ١)

| عدد الحجيج من خارج المملكة العربية السعودية | العام بالهجري | عدد الحجيج من خارج المملكة العربية السعودية | العام بالهجري |
|---|---|---|---|
| 1707814 | 1428 | 1431012 | 1423 |
| 1729841 | 1429 | 1419706 | 1424 |
| 1613965 | 1430 | 1534769 | 1425 |
| 1799601 | 1431 | 1557447 | 1426 |
| 1828195 | 1432 | 1654407 | 1427 |

**مجموعة البيانات (٣, ١ .ج) (بيانات مجتمع على شكل قيم عددية مرتبطة بمسمّيات):**

لتكن لدينا البيانات الآتية التي تبيّن لنا عدد سكان مناطق المملكة العربية السعودية (المصدر: وزارة الاقتصاد والتخطيط، مصلحة الإحصاءات العامة).

الجدول (٢ , ١)

| المنطقة | تعداد السكان لعام 2010 | المنطقة | تعداد السكان لعام 2010 |
|---|---|---|---|
| منطقة مكة المكرمة | 6915006 | منطقة تبوك | 791535 |
| منطقة الرياض | 6777146 | منطقة حائل | 597144 |
| المنطقة الشرقية | 4105780 | منطقة نجران | 505652 |
| منطقة عسير | 1913392 | منطقة الجوف | 440009 |
| منطقة المدينة المنورة | 1777933 | منطقة الباحة | 411888 |
| منطقة جازان | 1365110 | منطقة الحدود الشمالية | 320524 |
| منطقة القصيم | 1215858 | ----------- | ------- |

إنَّ مجموعات البيانات السّابقة تُعرف باسم **البيانات الخام** Raw Data، والدراسة هنا تتركّز بشكل أساسي على قيم وسلوك البيانات وليس على المسمّيات، ودور المسمّيات هنا للاستدلال بها والمقارنة فيما بينها، وأمّا الأزمنة فقد تكون من باب الاستدلال بها أو أن يكون لها دور مهمّ في الدراسة الإحصائية كما هو الحال في المتسلسلات الزمنية (انظر الفصل الثاني عشر).

الآن بفرض أنَّ جميع البيانات المطلوبة جُمعت دون نقصان، فإن كانت البيانات قليلة (وهنا عبارتي قليل وكثير تخضع لإمكانيات

الباحث وتقديراته، فهي عبارات نسبية ) فحينئذ يمكن استنتاج بعض التصورات حول سلوك هذه البيانات بالملاحظة، ومن ثمَّ يمكن للمرء الانتقال للتعامل مع المقاييس العدديّة للبيانات (سنأتي على ذكرها لاحقاً في الفصل الثاني) لاستنباط بعض المميّزات العددية لهذه البيانات. أمّا إذا كان عدد البيانات كبيراً فعندئذ عمليّة عرض جميع هذه البيانات واحدة فواحدة قد يكون مجهداً أو غير مجدٍ زمنيًّا أو ماديًّا، ولذلك يقوم المرء عادةً بعرض هذه البيانات وفق شكل يسمح له بأخذ انطباع أوليّ سريع عن هذه البيانات بأقل جهد ممكن، ولهذا السبب ابتكرت طرائق عديدة لعرض البيانات الإحصائية. إنَّ هذه الطرائق تتمايز بعضها عن البعض الآخر من حيث المبدأ والمنهج المعتمد في إنشائها، فمنها ما يُقدِّم البيانات في جداول رقمية ومنها ما يَعْرضها على شكل شرائط عمودية أو أفقيّة والآخر على شكل خطوط منكسرة أو منحنيّة و...، ولتوضيح ذلك سنقوم بشرح موجز لبعض الطرائق المستخدمة في عرض البيانات الإحصائية.

### (١ , ٣ , ١) طريقة الجدول Table Method

إنَّ طريقة الجدول لعرض البيانات تتلخص في أن يقوم المرء بتحميل (أو صبّ) البيانات التي لديه في جدول بتقسيمات معيَّنة حسب متطلبات المسألة المدروسة، وذلك بغية أخذ انطباع سريع عن سلوك البيانات وتسهيل العمليات الحسابيّة المتعلّقة بها، فعلى سبيل المثال نجد أنَّ مجموعتي البيانات رقم (٣ , ١ .ب) و(٣ , ١ .ج) السّابقتين مقدّمة جدولياً، وأمّا من أجل البيانات الخام رقم (٣ , ١ .أ) فيمكن عرضها في جدول بحيث يتمّ تدوين القيم في عمود، وفي العمود الآخر تدوَّن التكرارات المقابلة لتلك القيم (عدد المرّات التي تتكرّر فيها تلك القيم)، وهذا النوع من الجداول النّاتجة يُدعى جدولاً تكراريًّا بسيطًا (أو جدولًا تكراريًّا وذلك على سبيل التبسيط والاختصار)، ويُلجأ عادةً إلى هذا النوع من الجداول عندما يكون عدد البيانات الخام كبيراً. في الحقيقة توجد نماذج مختلفة من العرض الجدولي للبيانات ومنها:

#### (١ , ١ , ٣ , ١) الجداول التّكرارية Frequency Tables

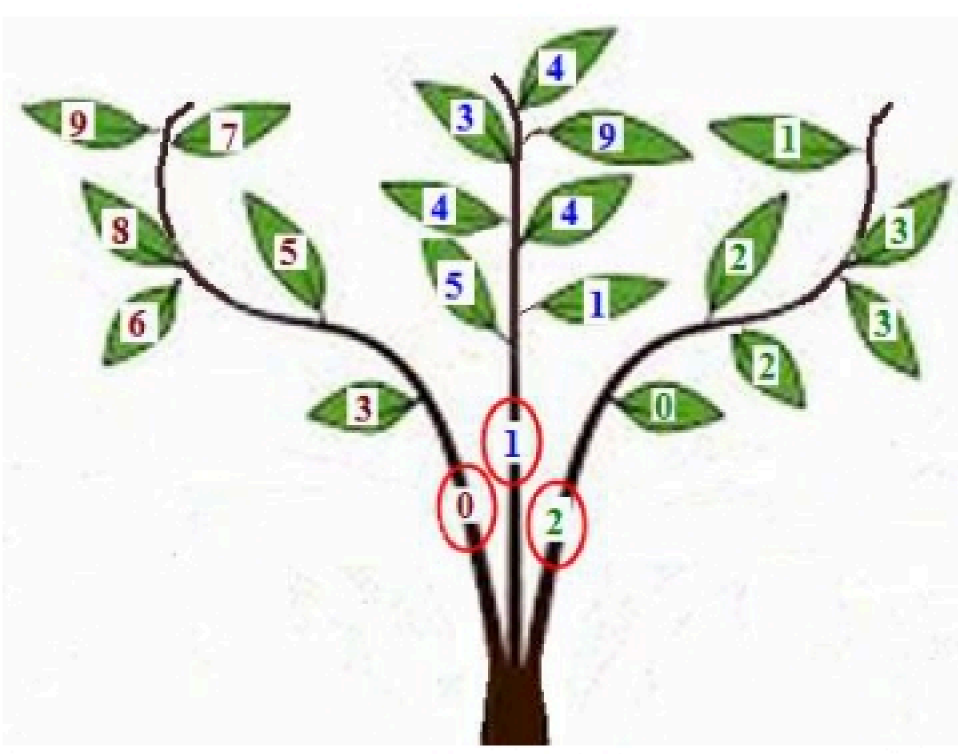
إنَّ الجداول التّكرارية تستخدم بكثرة في التجارب المعمليّة ووسائل الإعلام لسهولة بنائها وفهمها من قبل الكثير من الأشخاص، ويُبنى الجدول التكراري من خلال صبّ البيانات الخام في جدول بحيث يتمّ تدوين قيم الممثِّلين للبيانات في عمود (أو صف) (فعلى سبيل المثال في الجدول الآتي لدينا القيمة $x_3 = 5$ ممثِّل لكل البيانات التي لها القيمة 5)، وفي العمود الآخر (أو الصف الآخر) تدوَّن التكرارات المقابلة لقيم ممثِّلي البيانات (في الجدول الآتي نجد لممثِّل القيمة 5 تكرار يساوي 3، وهذا يعني أنّه لدينا ثلاثة بيانات لها القيمة 5)، ومن ثمّ يمكن عرض مجموعة البيانات (٣ , ١ .أ) المُعطاة في جدول تكراري وفقاً للعرض الآتي:

الجدول (٣ , ١) خاص بمجموعة البيانات (٣ , ١ .أ)

| $i$ | عمر البطارية (بالشهر) $x_i$ | التكرار (عدد البطاريات) $f_i$ Frequency of $x_i$ | $i$ | عمر البطارية (بالشهر) $x_i$ | التكرار (عدد البطاريات) $f_i$ Frequency of $x_i$ |
|---|---|---|---|---|---|
| 1 | 2 | 1 | 7 | 11 | 10 |
| 2 | 4 | 2 | 8 | 12 | 12 |
| 3 | 5 | 3 | 9 | 13 | 9 |
| 4 | 6 | 4 | 10 | 15 | 7 |
| 5 | 7 | 5 | 11 | 17 | 3 |
| 6 | 9 | 7 | 12 | 19 | 1 |
| Total | ----------- | ---------------- | ----- | ---------------- | 64 |

### (١,٣,١,٢) جدول الساق والأوراق Stem and Leafs Table

توجد طريقة أخرى لصبِّ البيانات الخام في جدول مُشابه للجداول التّكرارية تُدعى **عرض الساق والأوراق**، والغاية منها تصنيف البيانات في مجموعات جزئية، كأن نبيّن القيم ما بين 10 و19 في مجوعة جزئية واحدة، والقيم ما بين 20 و29 في مجوعة جزئية ثانية، والقيم ما بين 70 و79 في مجوعة جزئية أخيرة، وقد يُعتمد أسلوب آخر للتصنيف ولكن وفقاً لهذه الآلية. بمعنى آخر، إنَّ جدول السـاق والأوراق هو عرض جدولي توضيحي يبيِّن تكرار حدوث قيم معينة، ويتمّ ذلك من خلال صب البيانات الخام في جدول بعمودين أحدهما يخصَّص للساق Stem والآخر للأوراق Leafs، وذلك على النحو الآتي:




بفرض أنَّه لدينا بيانات على شكل أعداد صحيحة مكوَّنة من خانتين على الأكثر (كما في المثال الآتي وذلك على سبيل التبسيط في الشرح)، فعندئذ نضع في عمود الساق القيمة 0، وفي العمود المقابل لها (عمود الأوراق) نضع جميع القيم المكوَّنة من خانة واحدة فقط، وبعد ذلك ننتقل إلى الأرقام المكوَّنة من خانتين فنبدأ بالقيم التي خانة العشرات لها هي الأصغر، فنضع العدد المكوِّن لخانة العشرات في عمود الساق وأمّا الأجزاء الأخرى المكوِّنة لخانة الآحاد من العدد فنضعها في عمود الأوراق، فإذا ما انتهينا من هذه ننتقل للرقم الأكبر والمكوَّن من خانتين أيضاً وهكذا دواليك حتى الانتهاء من عملية الصبّ، ويفضَّل عادةً ترتيب الأوراق تصاعدياً.

#### (١,٣,١,٢,١) مثال

لتكن البيانات الآتية التي تعرض عدد ساعات العمل اليوميّة لحاسب مركزي في مؤسسة تعليمية:

15 3 14 13 14 11 21 5 14 19 9 21 20 6 8 22 23 23 7

عندئذلصب هذه البيانات في جدول الساق والأوراق سنقوم أولاً بوضع صفر إلى يسار الأعداد التي أصغر من عشرة (وذلك على سبيل التبسِّط)، فيصبح للبيانات الُمعطاة العرض الآتي:

15 03 14 13 14 11 21 05 14 19 09 21 20 06 08 22 23 23 07

ولدى صبِّ البيانات الُمعطاة في جدول نحصل على الجدول الآتي.

الجدول (٤,١.أ) الأوراق بحسب موضعها في البيانات المقدَّمة

| الساق Stem | الأوراق Leafs | التكرار Frequency |
|---|---|---|
| 0 | 3, 5, 9, 6, 8, 7 | 6 |
| 1 | 5, 4, 3, 4, 1, 4, 9 | 7 |
| 2 | 1, 0, 2, 2, 3, 3 | 6 |
| Total | ------------- | 19 |

الجدول (٤ , ١ .ب) الأوراق مرتّبة تصاعدياً

| الساق Stem | الأوراق Leafs | التكرار Frequency |
|---|---|---|
| 0 | 3, 5, 6, 7, 8, 9 | 6 |
| 1 | 1, 3, 4, 4, 4, 5, 9 | 7 |
| 2 | 0, 1, 2, 2, 3, 3 | 6 |
| Total | ------------- | 19 |

الآن، وفي حال كانت البيانات مكوَّنةً من أعداد غير صحيحة (جزء صحيح وآخر عشري)، فعندئذ نضع في عمود الساق أصغر قيمة صحيحة مع قيمة الجزء العشري الأصغر قيمةً، وفي العمود المقابل لها (عمود الأوراق) نضع جميعَ القيم المتبقيَّة والمكوَّنة للجزء العشري بعد الرقم الذي دوِّن في عمود الساق. بعد ذلك ننتقل إلى الأرقام المكوّنة القسم الصحيح نفسه (إن وجد) ولكن بجزء عشري أول أكبر من سابقه ونجري على هذه الأرقام التصنيف نفسه الذي تمَّ على الأرقام السابقة، فإن انتهينا من الأجزاء العشرية ننتقل للجزء الصحيح الأكبر من السابق، وهكذا دواليك حتى الانتهاء من عملية الصب، والمثال الآتي يوضِّح لنا ذلك.

**مثال (٢ , ٢ , ١ , ٣ , ١)**

لتكن البيانات الآتية التي تعرض أوزان عشرين حبة قرع في مزرعة (مقدَّرة بالكيلو غرام):

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.14 | 0.13 | 0.12 | 1 | 1.01 | 1.05 | 1.11 | 1.12 | 1.3 | 1.3 | 1.3 | 1.32 | 1.39 |
| 1.34 | 1.44 | 2.23 | 2.29 | 2.24 | 2.2 | 3.19 | 3.15 | 3.12 | 3.17 | 3.60 | 3 | |

فعندئذ لصب هذه البيانات في جدول الساق والأوراق يُفضّل (وعلى سبيل التبسيط) القيام أولاً بإضافة صفرين بعد الفاصلة إلى يمين كل عدَد صحيح (كأن نكتب 1 بالشكل 1.00)، وصفر بعد الرقم العشري الأول بعد الفاصلة للأرقام ذات المرتبة العشرية الواحدة (كأن نكتب 1.3 بالشكل 1.30)، وذلك لجعل المراتب العشرية متماثلة، فيصبح للبيانات المُعطاة العرض الآتي (نشير هنا إلى أنَّه في بعض الأحيان يقدَّم الجدول بعد ترتيب بيانات الأوراق تصاعديًّا أيضاً):

الجدول (٥ , ١)

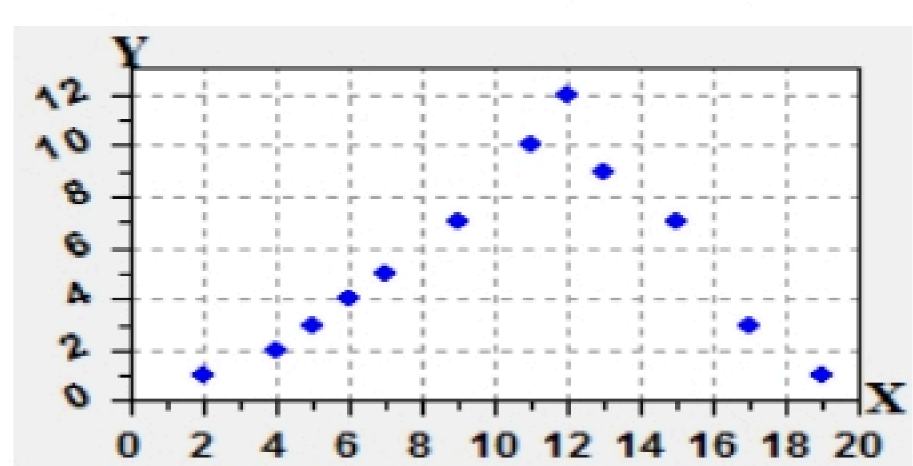
| الساق Stem | الأوراق Leafs | الأوراق بعد ترتيبها تصاعدياً Leafs | التكرار Frequency |
|---|---|---|---|
| 0.1 | 3, 4, 2 | 2, 3, 4 | 3 |
| 1.0 | 1, 0, 5 | 0, 1, 5 | 3 |
| 1.1 | 1, 2 | 1, 2 | 2 |
| 1.3 | 0, 0, 0, 2, 9, 4 | 0, 0, 0, 2, 4, 9 | 6 |
| 1.4 | 4 | 4 | 1 |
| 2.2 | 9, 3, 4, 0 | 0, 3, 4, 9 | 4 |
| 3.0 | 0 | 0 | 1 |
| 3.1 | 9, 5, 2, 7 | 2, 5, 7, 9 | 4 |
| 3.6 | 0 | 0 | 1 |
| Total | ------------- | ------------- | 25 |

كما يوجد نوع آخر من التمثيلات الجدولية للبيانات يسمَّى جدول التوزيع التكراري حيث سنأتي على شرحه لاحقاً في هذا الفصل.

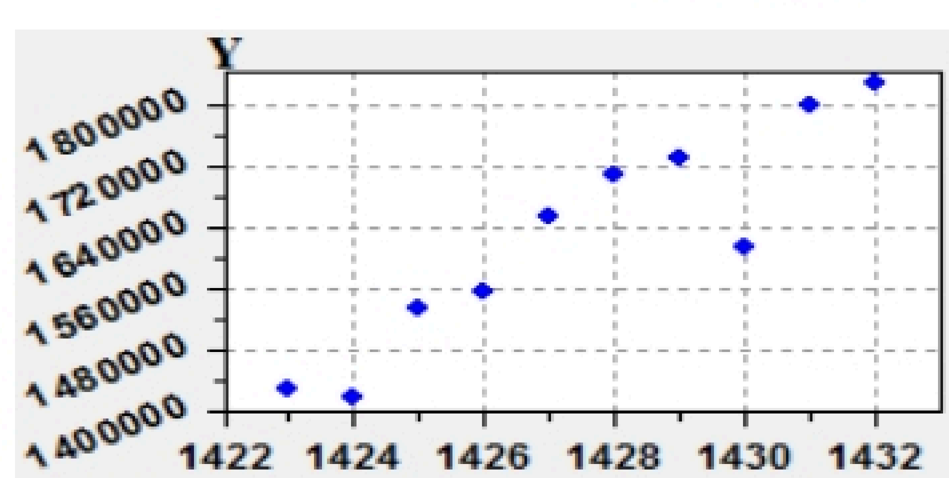
### (١,٣,٢) التمثيل باستخدام لوحة الانتشار Scatter Plot

إنَّ لوحة الانتشار هو تلك اللوحة المتمثِّلة بالمستوى الإحداثي **XoY** لدى استخدام تجزئات مُحَدَّدة لكل من المحورين **oX** و **oY** بحسب الهدف المنشود. أمّا عملية تمثيل البيانات الخام عليها فإنَّها تتلخص بوضع القيم (أو المسمَّى أو السّنة أو الفترة الزمنية) أفقياً، ومن ثَمَّ رسم نقطة في المستوى الإحداثي **XoY** بحيث تمثِّل القيم (أو المسمَّى أو السنة أو الفترة الزمنية) على المحور **oX** في حين يدوَّن على المحور **oY** القيم (أو التكرارات) المقابلة لتلك القيمة (أو المسمَّى أو السنة أو الفترة الزمنية) وذلك باستعمال مقياس رسم مناسب، وتدعى بالتمثيل النقطي Dots Representation للبيانات أيضاً. إنَّ هذه الطريقة تستخدم للمقارنة بين قيم الظواهر حسب الزمن أو المسمَّيات أو لمقارنة ظواهر بعضها مع بعض.

تجدر الإشارة هنا إلى أنَّ العرض البياني للنقاط الممثِّلة للبيانات على المستوى الإحداثي **XoY** يُدعى **العرض الانتشاري** للبيانات، فعلى سبيل المثال نجد العرض الانتشاري لمجموعتي البيانات (١,٣.أ) و (١,٣.ب) كما في الشّكلين الآتيين:

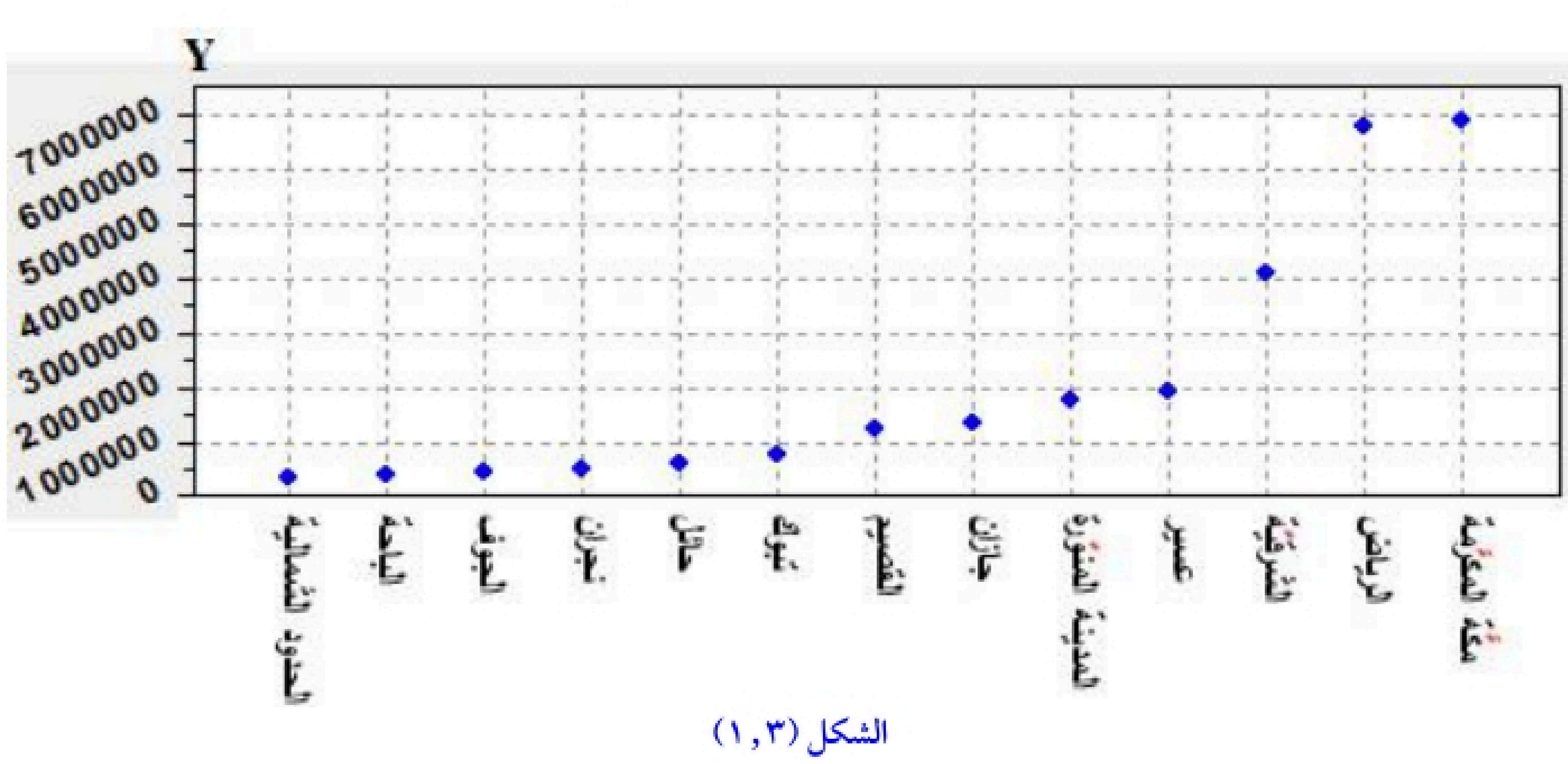



الشكل (١,١) أعمار بطاريات أجهزة حاسب محمول.




الشكل (١,٢) أعداد الحجيج من خارج المملكة.

وأخيراً نجد العرض الانتشاري لمجموعة البيانات (١,٣.ج) كما في الشّكل الآتي:




الشكل (١,٣)

### (١,٣,٣) العرض الشرائطي Bar Graph

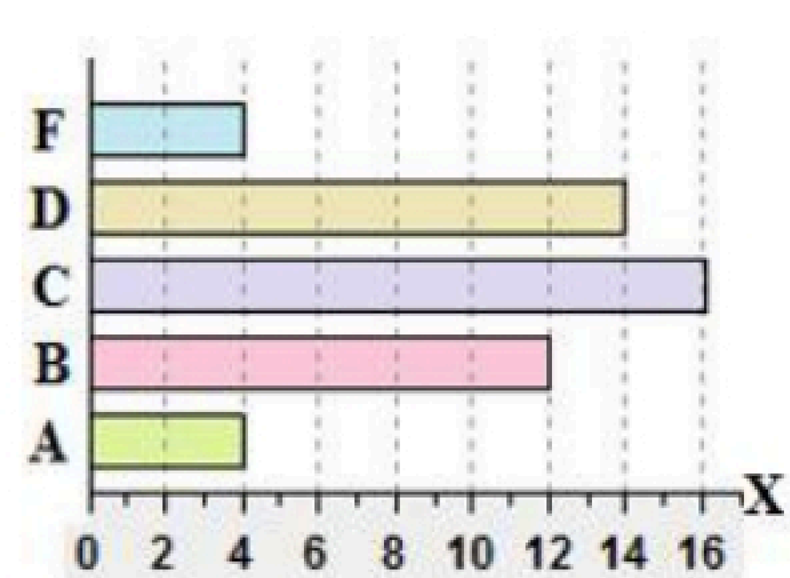
تقوم طريقة العرض الشرائطي للبيانات على تمثيل البيانات من خلال شرائط مستطيلة رأسيّة أو أفقيّة (وفي العروض ثلاثية الأبعاد قد تستخدم نماذج أخرى كالأعمدة الأسطوانية أيضاً) منفصلة بعضها عن بعض، وتتلخص هذه الطريقة بوضع المسمَّيات أو الزمن أفقياً أو رأسيّاً ومن ثمَّ رسم شريط مستطيل (عموماً إذا كان رأسياً أو شريط إذا كان أفقياً) على كل مسمًّى (أو كل سنة أو كل فترة زمنية) بحيث يكون طول كل شريط مُمثِّلاً للقيمة (أو التكرار) المقابل لذلك المسمَّى (أو لتلك السنة أو الفترة الزمنية) وذلك باستعمال مقياس رسم مناسب، وتُستخدم هذه الطريقة للمقارنة بين قيم الظواهر حسب المسمَّيات أو الزمن، أو لمقارنة ظواهر بعضها مع بعض.

#### (١,٣,٣,١) أمثلة

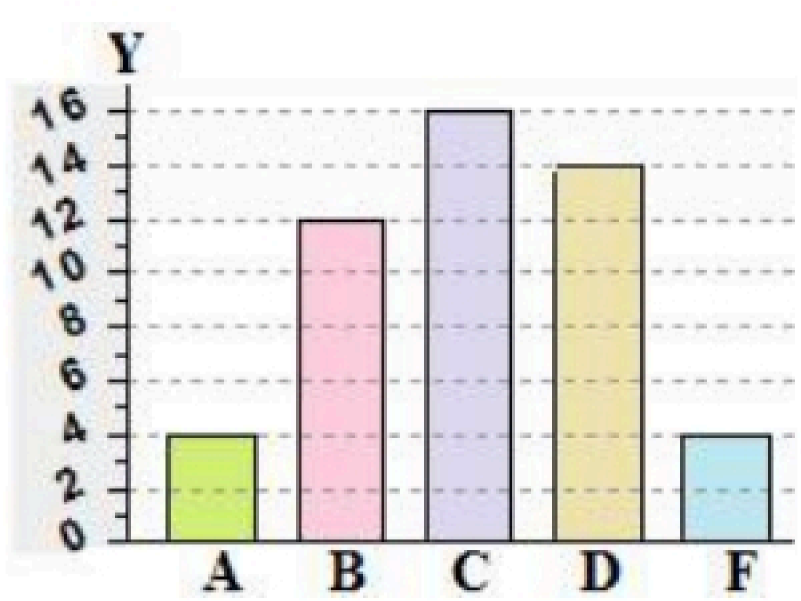
١- البيانات الآتية تمثِّل تقديرات 50 طالباً في مقرَّر الإحصاء:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| F | C | D | C | A | B | C | F | D | D |
| C | D | C | C | D | A | B | C | D | B |
| D | C | F | D | C | B | B | A | D | C |
| B | D | C | B | B | C | B | C | D | C |
| B | A | D | C | B | D | C | D | B | F |

فنجد العرض الشرائطي الأفقي والرأسي لهذه البيانات كما في الشّكلين الآتيين:

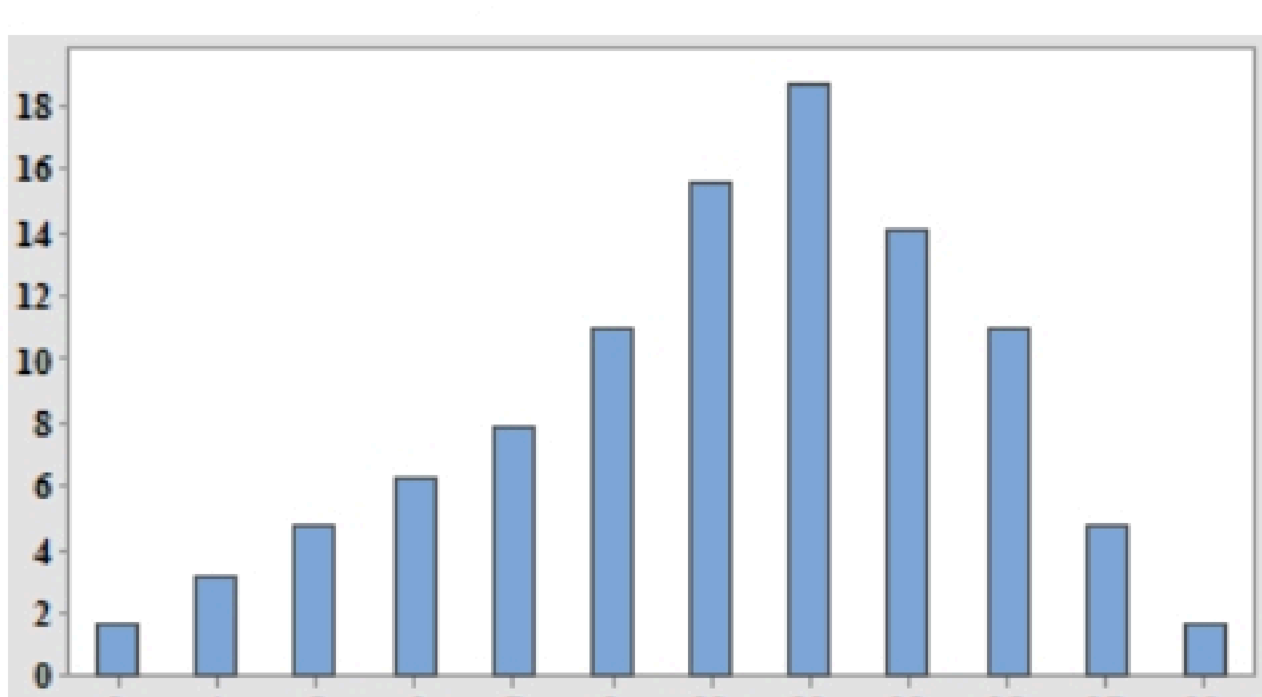



الشكل (٤,١.أ) العرض بالشرائط الأفقيّة.



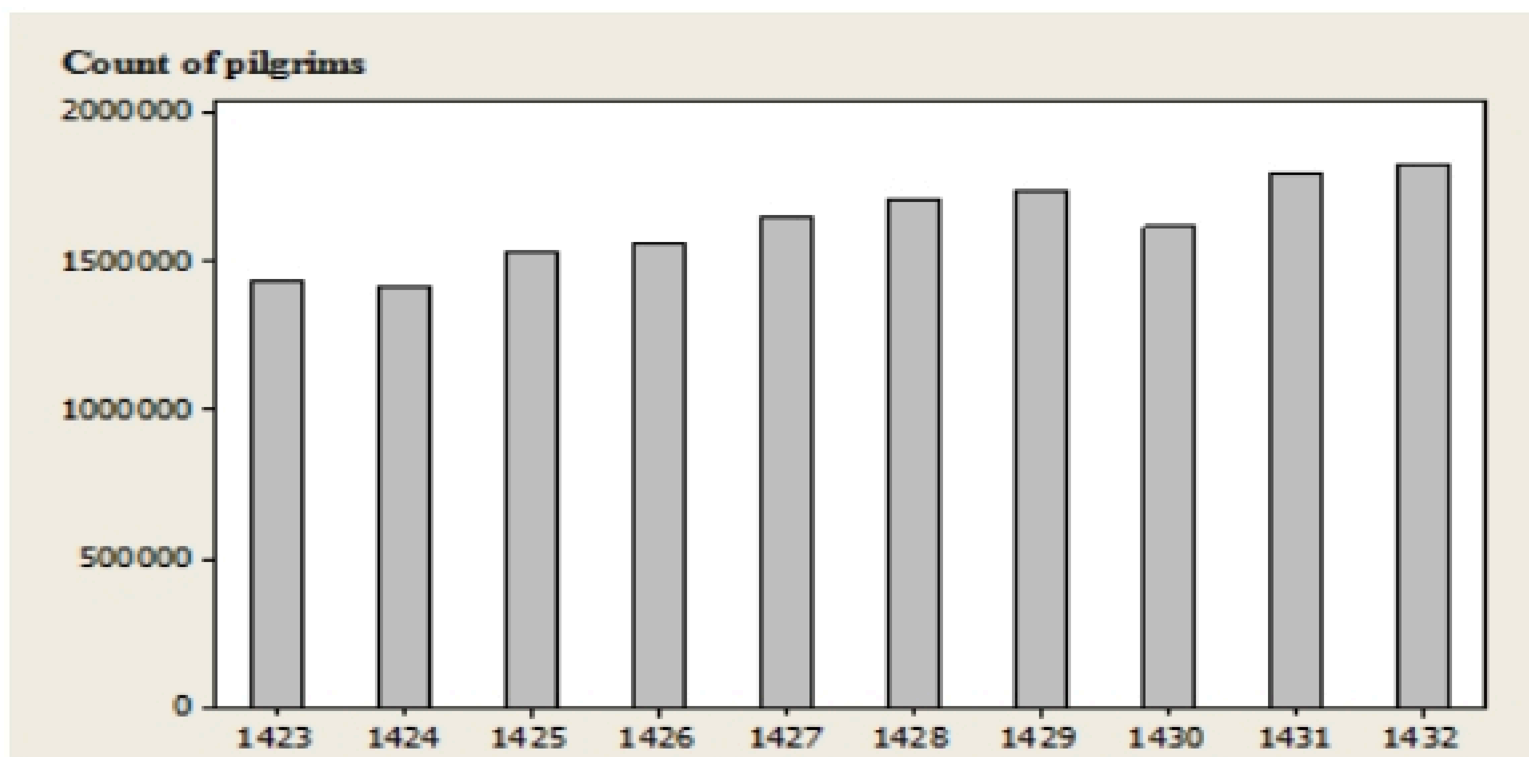

الشكل (٤,١.ب) العرض بالشرائط الرأسيّة (بالأعمدة).

٢- بالعودة إلى مجموعة البيانات (٣,١.أ) نجد أنَّ لعرضها الشرائطي الشكل الآتي مستخدمين الأعمدة في ذلك:



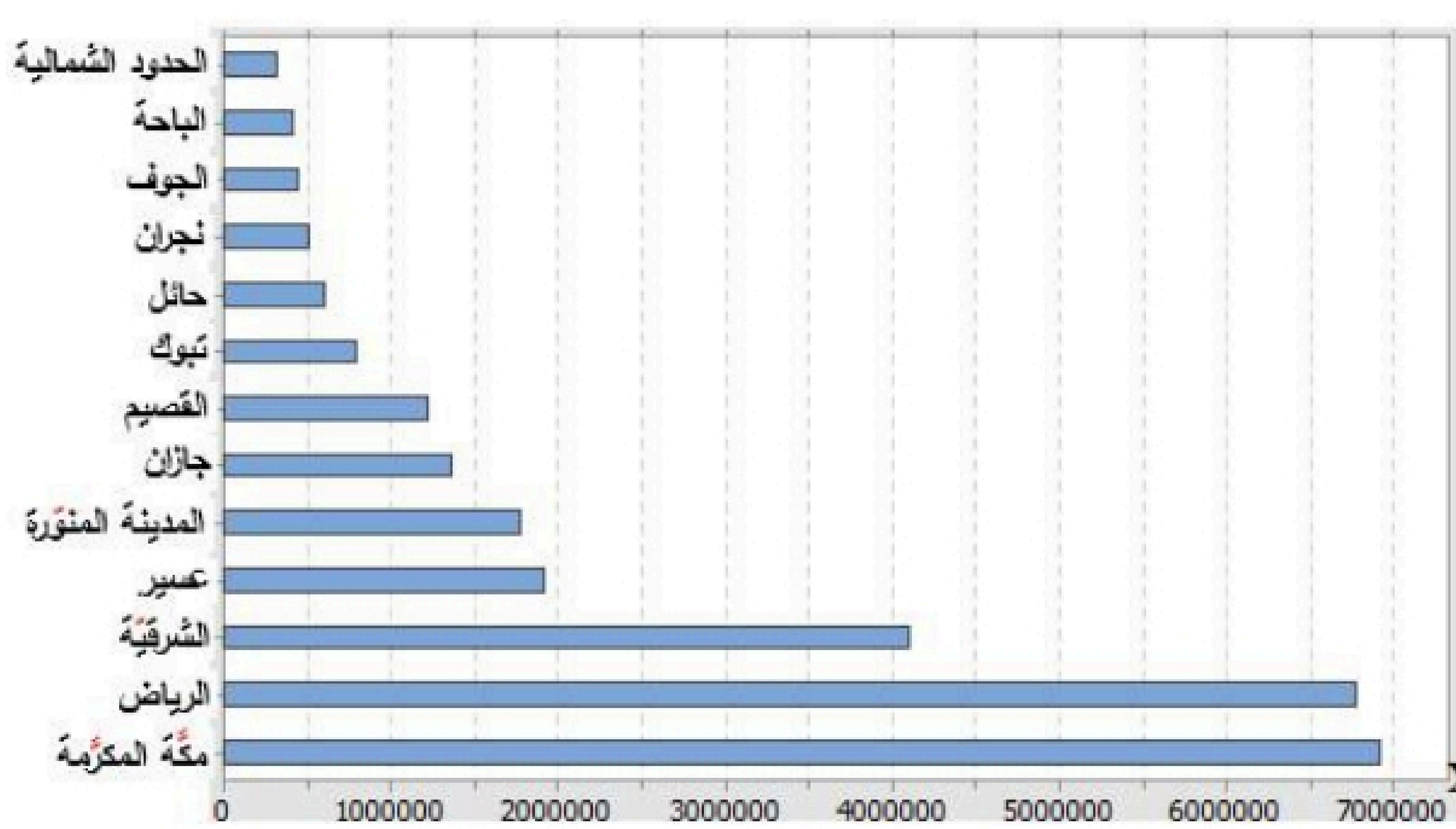

الشكل (١,٥)

٣- بالعودة إلى مجموعة البيانات (٣,١.ب) نجد أنَّ لعرضها الشرائطي الشكل الآتي مستخدمين الأعمدة في ذلك:




الشكل (١,٦) أعداد الحجيج من خارج المملكة.

٤- بالعودة إلى مجموعة البيانات (٣,١.ج) نجد أنَّ لعرضها الشرائطي الشكل الآتي مستخدمين في ذلك الأشرطة الأفقيّة (لاحظ هنا أنَّه قد يكون التمثيل بالأشرطة الأفقيّة أكثر ملاءمةً من الأشرطة الرأسيّة لوجود مسميّات):




الشكل (١,٧) تعداد السكان لمناطق المملكة العربيّة السعوديّة.

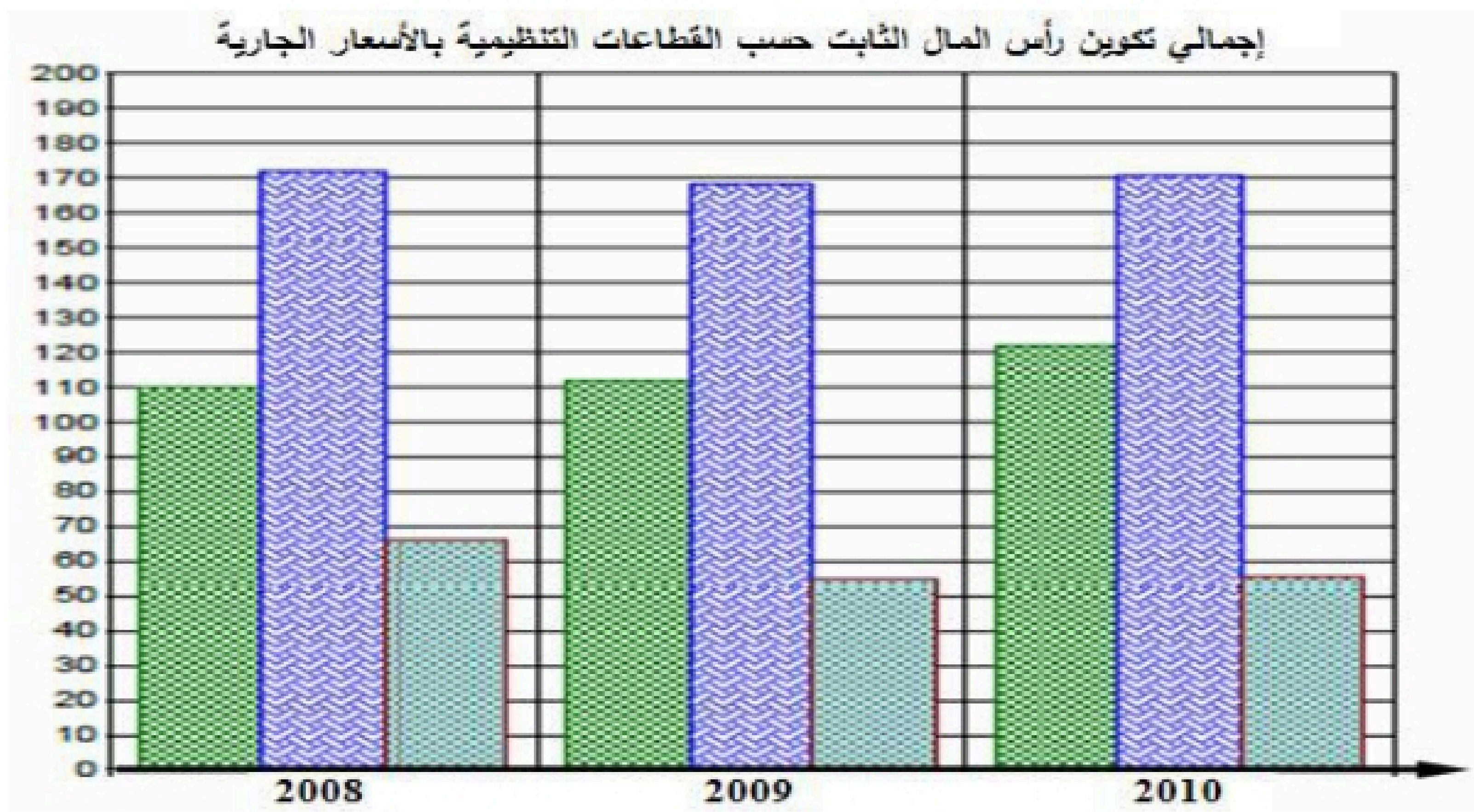
### (١,٣,٣,٢) ملاحظات

١- من الممكن أن يكون لدينا ممثلين (كالمسميات مثلاً) للبيانات ويقابها قيم موجبة وأخرى سالبة (مثل درجات الحرارة)، فعندئذ سيكون لدينا عروضاً شرائطية تتَّجه باتجاهين متضادين حيث تدعى أمثال هذه العروض بـ **العروض الشرائطية ثنائية الاتجاه**.

٢- يمكن تقديم العروض الشرائطية لمجموعتين أو أكثر من البيانات بآن واحد (في عرض واحد) وذلك لجعل المقارنة بين العروض البيانية أكثر يسراً وسهولةً، ومن هذه العروض:

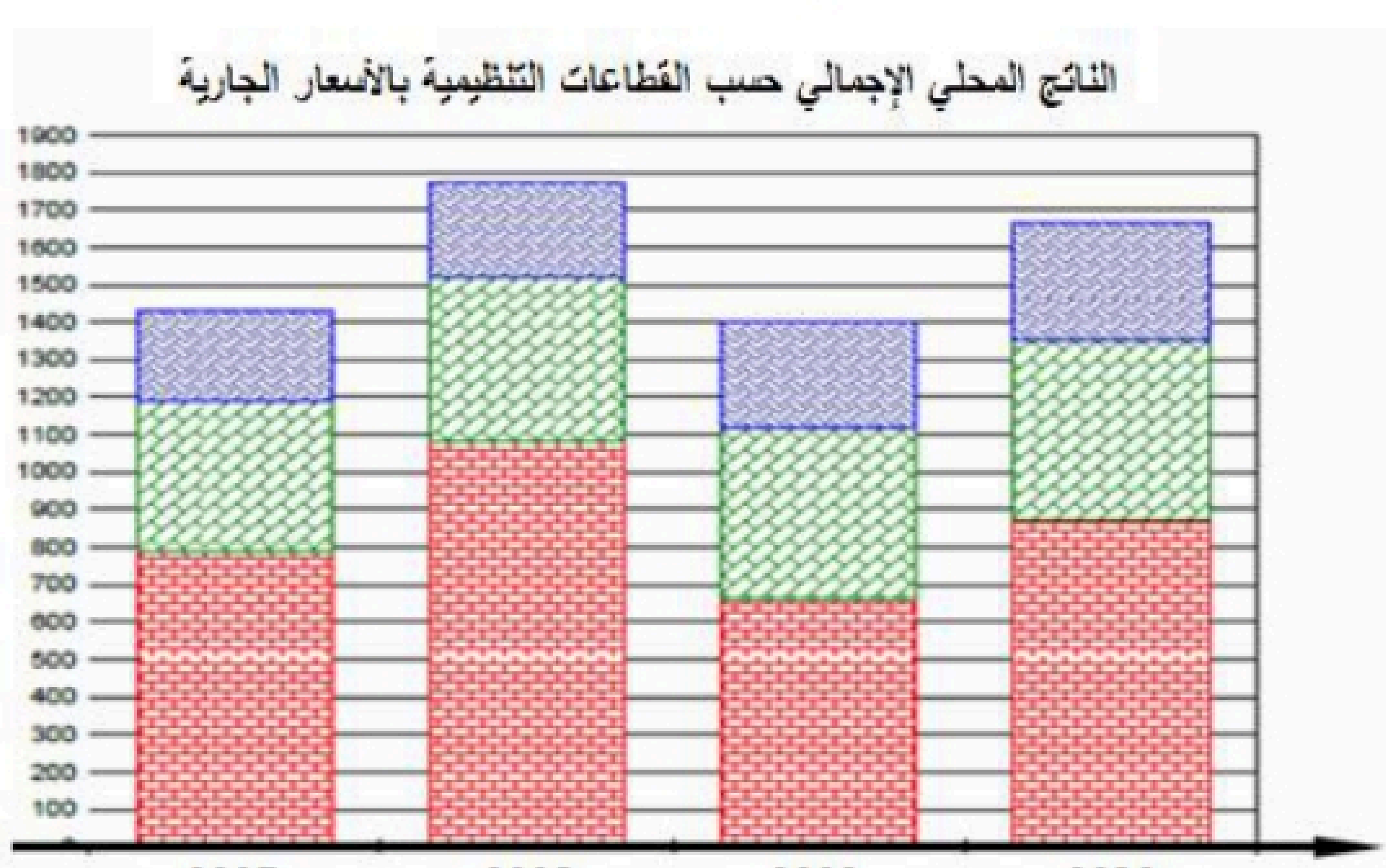
أ- **طريقة الشرائط المتجاورة** Contiguous Bar Charts، وتستخدم هذه الطريقة أسلوبًا مماثلاً تماماً للتمثيل الشرائطي من أجل كلّ مجموعة بيانات على حدة، وبعد ذلك توضع الأشرطة المثّلة للشيء الواحد قريبة أو ملاصق بعضها لبعض، ويميَّز أحدهما

عن الآخر بالتلوين المختلف أو التظليل المختلف، والشكل الآتي يقدِّم لنا نموذج لهذا العرض:




الشكل (١,٨)

ب- **طريقة الشرائط المتركبة** Overlapped Bar Charts، وتستخدم هذه الطريقة أسلوبًا مماثلاً تماماً للتمثيل الشرائطي من أجل كلّ مجموعة بيانات على حدة أيضاً، ولكن توضع الأشرطة الممثِّلة للشيء الواحد بعضها فوق بعض، ويميِّز أحدهما عن الآخر بالتلوين المختلف أو التظليل المختلف، والشكل الآتي يقدِّم لنا نموذجًا لهذا العرض (المصدر مصلحة الإحصاءات العامة والمعلومات، الحسابات القومية 1431-1432 هـ، 2010 م):



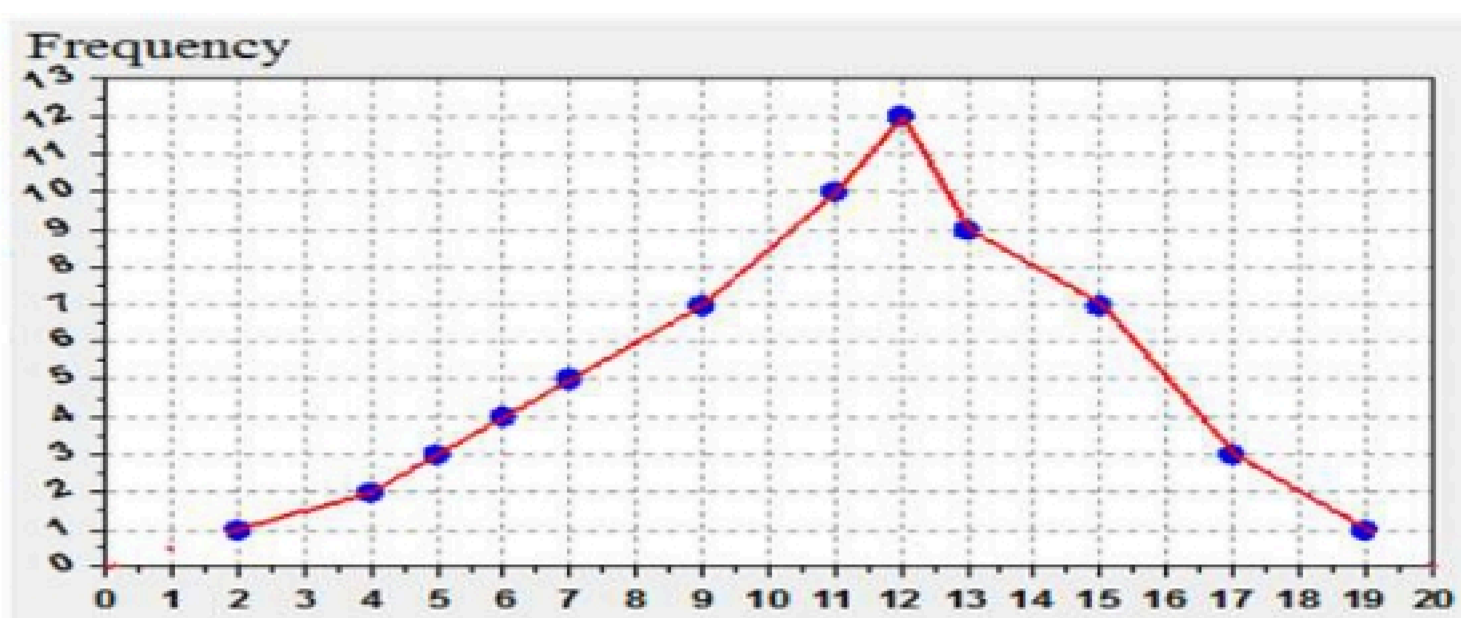

الشكل (١,٩)

### (١,٣,٤) المضلَّع Polygon

تستخدم هذه الطريقة في عرض تَغيُّر ظاهرة (أو عدة ظواهر) مع الزمن أو مسمًّى لشيء ما، ومن الأمثلة على ذلك:

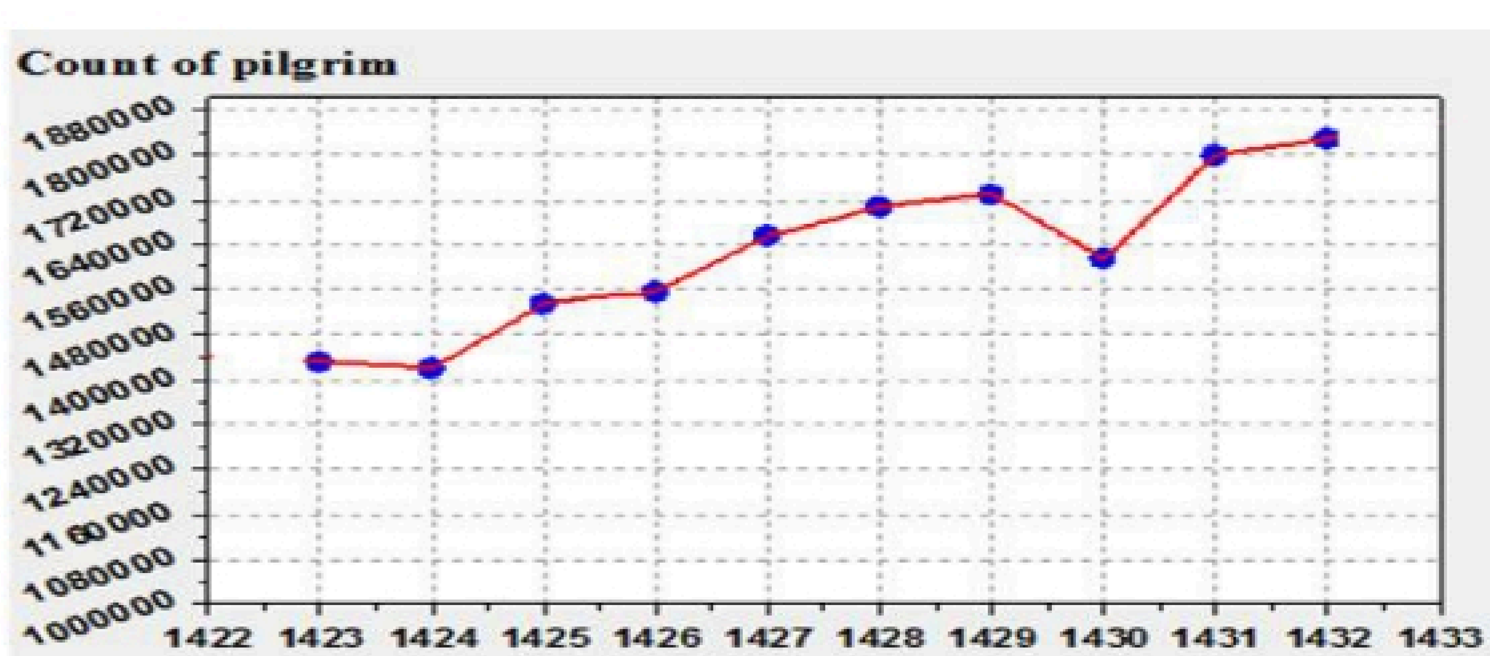
- تَغيُّر درجة الحرارة في مدينة ما مع مرور الزمن.
- تَغيُّر نسبة السكر في نبات الشمندر السكري لدى أنواع مختلفة منه.
- تَغيُّر الطول لأنواع مختلفة من المعادن لدى تعرضها لدرجة حرارة معيَّنة.

إنَّ التمثيل بخطوط مستقيمة متتالية تكوِّن بمجملها خطًّا منكسًرا يتمّ برسم محورين متعامدين، ومن ثَّم رصد قيم الظاهرة كنقاط ذات إحداثيين بحيث يُمثِّل المحور الإحداثي **oX** الزمن أو المسمَّيات في حين يُمثِّل المحور الإحداثي **oY** قيم المتغيِّر، وبعد تمثيل كل القيم بنقاط نقوم بوصل كل نقطتين متتاليتين بقطعة مستقيمة فنحصل بذلك على خط منكسر يُمثِّل تَغيُّر الظاهرة المدروسة مع الزمن (أو المسمَّيات أو...)، وهذا التمثيل يدعى مضلَّعًا، فعلى سبيل المثال لو أردنا عرض البيانات (٣,١.أ) بوساطة المضلَّع (أو الخطوط المنكسرة) فإنَّنا نجد لها التمثيل البياني الآتي:



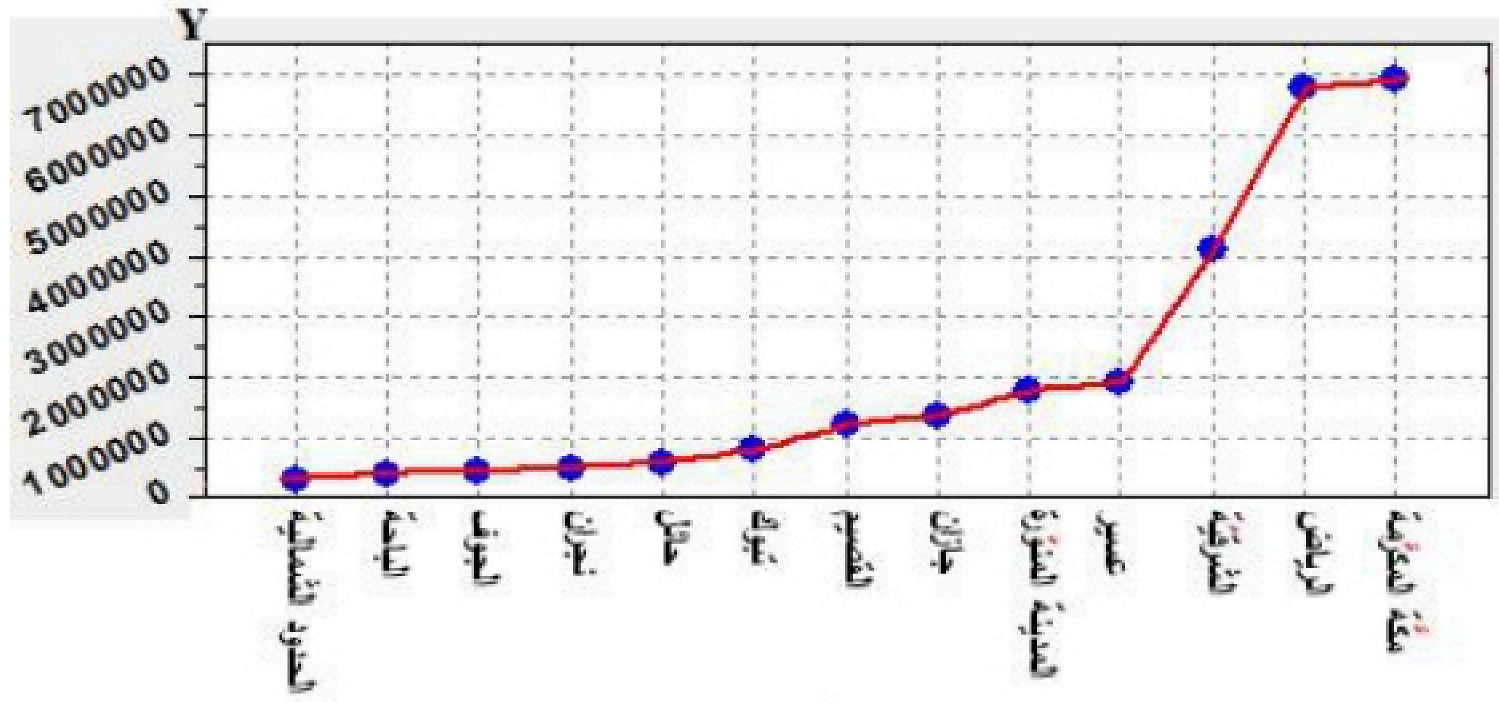

الشكل (١,١٠)

وأمّا البيانات (٣,١.ب) فنجد لها المضلَّع الآتي:




الشكل (١,١١)

وأخيراً نجد العرض الآتي للبيانات (٣,١.ج) لدى استخدام المضلَّع:




الشكل (١,١٢)

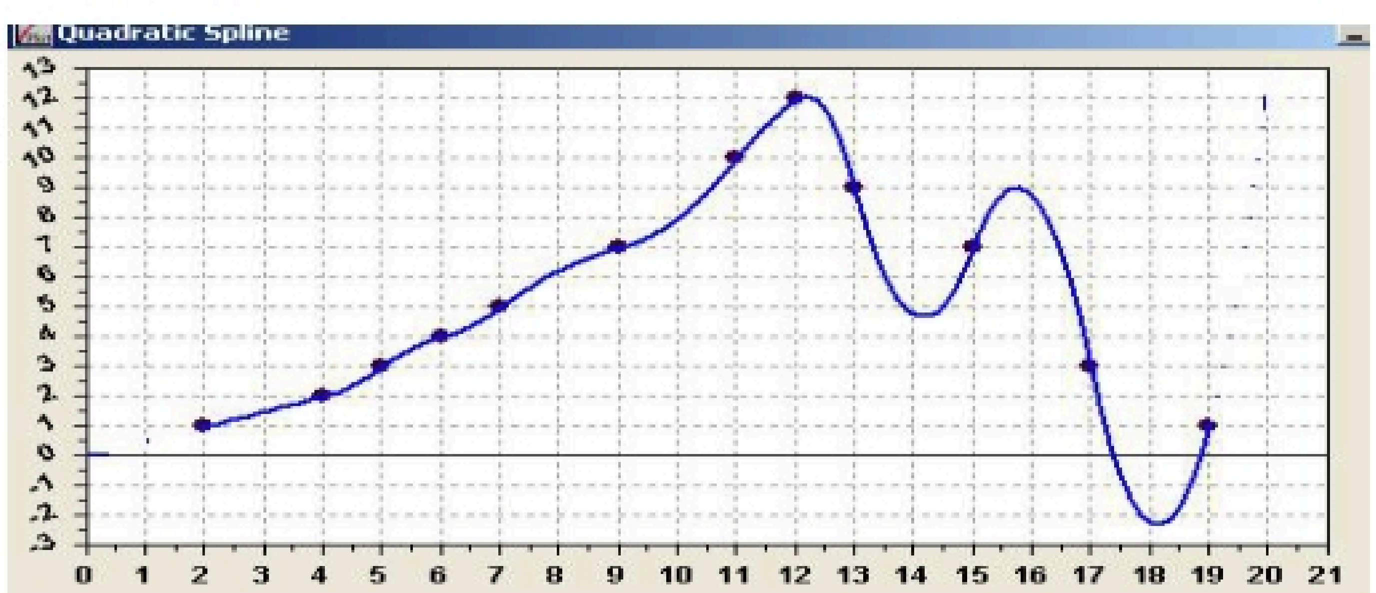
### (١,٣,٥) الخط المنحني Curve:

إنَّ هذه الطريقة تماثل طريقة العرض باستخدام المضلّعات (الخطوط المنكسرة) حيث نحصل عليه بتمهيد الخط المنكسر ليصبح منحنياً من خلال إحدى الطرائق الآتية:

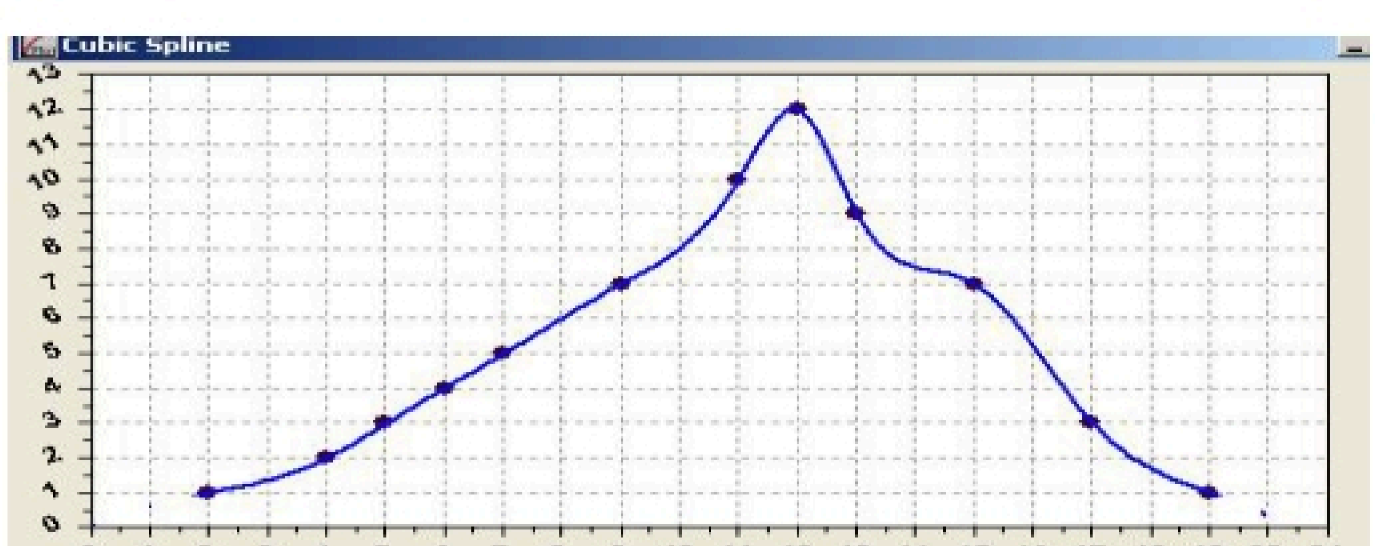
١- جعل الفترات التي تبنى عليها هذه الخطوط المنكسرة قصيرة جداً مع زيادة عددها بشكل كبير جداً، ولذلك فإنَّ هذه الطريقة نادرة الاستخدام في المجالات التطبيقية.

٢- نقوم بالتمهيد اليدوي للخطوط المستقيمة لتصبح منحنية وملائمةً عند مرورها بالنقاط الممثِّلة للبيانات، وهذه الطريقة قليلة الاستخدام أيضاً لأنها تحتاج إلى مهارات في الرسم.

٣- استخدام أحد البرامج في عمليّة التمهيد كأن نستخدم عملية **التمهيد الشرائحي** Spline من الدرجة الثانية أو الثالثة (متوفرة في برنامج Curve Expert) أيّهما أنسب للعرض وأقل تناقضاً مع واقع البيانات، وهذه الطريقة تعطينا نتائج ممتازة لعملية التمهيد، فعلى سبيل المثال لو أردنا عرض البيانات (٣,١.أ) من خلال التمهيد الشرائحي من الدرجة الثّانية والثّالثة فنجد لها التمثيلين البيانيين الآتيين على الترتيب:




الشكل (١٣,١.أ). تمهيد شرائحي من الدرجة الثانية.




الشكل (١٣,١.ب). تمهيد شرائحي من الدرجة الثالثة.

فنلاحظ أنَّ المنحنى الناتج عن التمهيد الشرائحي من الدرجة الثالثة أكثر ملائمةً لواقع البيانات، ولكن هذا لا يعني أنَّ رفع درجة التمهيد الشرائحي ستعطينا نتائج أفضل دوماً.

### (١,٣,٦) القرص الدائري Pie Chart

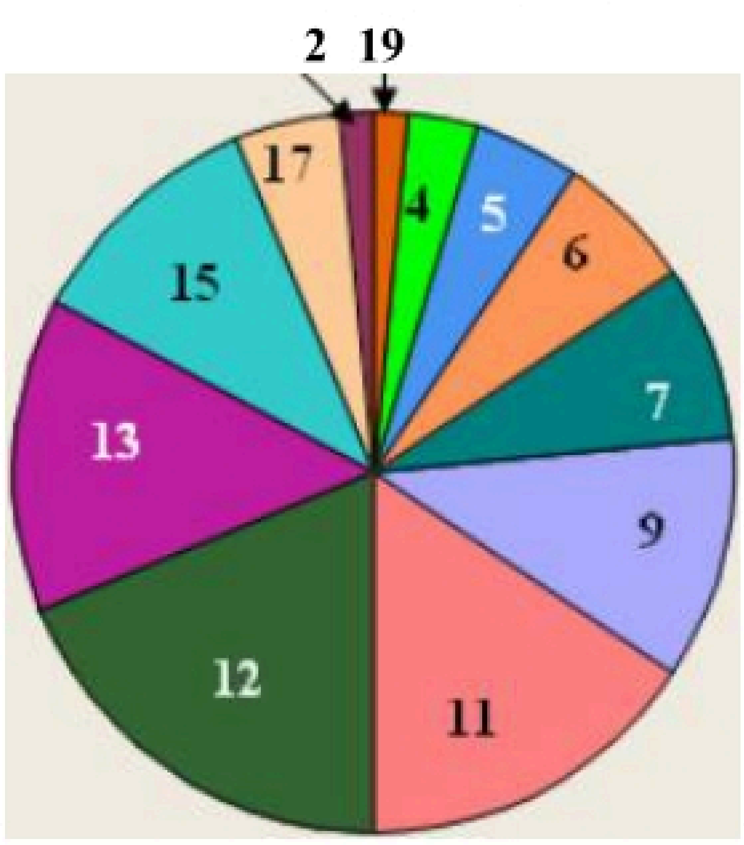
إنَّ تمثيل الرسوم البيانية باستخدام القرص شائع بكثرة وخاصّة في وسائل الإعلام، وهي تُعطي صورة مفيدة عن قيم النسبة المئوية، وخاصة عند مقارنة مجموعات مختلفة من البيانات، إلا أنَّ رسمها أكثر صعوبة من الحالات السابقة إذ تحتاج إلى فرجار ومنقلة، أو استخدام البرامج الإحصائية (وهي متوفرة لدى معظم البرامج الإحصائية). على أي حال، يُلجأ إلى استعمال هذه الطريقة عندما نكون

بحاجة لتقسيم الكل إلى $k$ من الأجزاء، وأمّا لرسمها فإنّنا نقوم أولاً برسم دائرة بنصف قطر مُثّبت (غالباً يكون عمودياً) يُعدّ مبدأً لقياس الزاوية عنه، ثمّ نحسب الزوايا $\alpha_i$ (مقدّرة بالدرجات Degrees) وتأخذ إلى يمين العمود السابق باتجاه دوران عقارب الساعة، وبحيث يكون للجزء $i$ قطاع دائري زاويته تُحسب بوساطة العلاقة الآتية:

$$\alpha_i := \frac{n_i}{n} \times 360 \text{ for a sample with size } n \qquad [1,1]$$

$$\alpha_i := \frac{n_i}{N} \times 360 \text{ for a population with size } N \qquad [1,2]$$

علماً أنّ $n_i$ هو عدد العناصر (أو البيانات) التابعة للجزء $i$ مع $i \in \mathbb{N}_k$، فعلى سبيل المثال لو أردنا عرض البيانات (١,٣.أ) من خلال تجزئة القرص (حيث لدينا حجم العيِّنة $n = 64$)، فإنّنا نجد التمثيل البياني لتلك البيانات له الشكل الآتي (١,١٤)، علماً أنّ قيمة زاوية القطاع الدائري الأول (الموافق للقيمة 2) تساوي:



الشكل (١,١٤)

الدرجات degrees

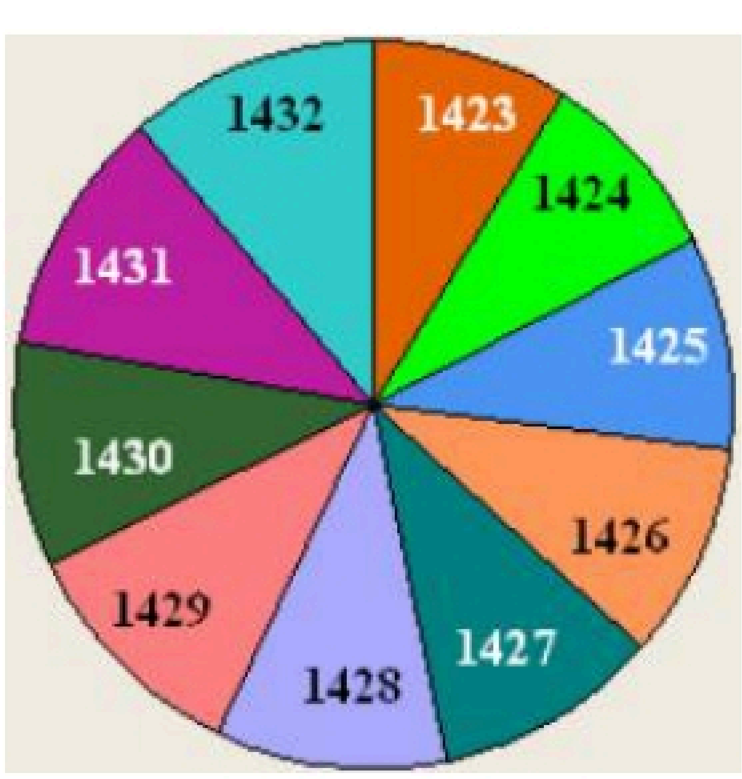
$$\alpha_1 = \frac{n_1}{n} \times 360 = \frac{1}{64} \times 360 = 5.63 = 5:37;48$$

الجزء العشري من الدرجة | الدقائق | الثواني

في حين نجد أنّ قيمة زاوية القطاع الدائري السادس (الموافق للقيمة 9) تساوي:

$$\alpha_6 = \frac{n_6}{n} \times 360 = \frac{7}{64} \times 360 = 39.38 = 39:22;48$$

أمّا إذا أردنا عرض البيانات /١,٣.ب/ من خلال تجزئة القرص (حيث لدينا حجم العيِّنة $n = 16276757$)، فإنّنا نجد لها التمثيل البياني المقدَّم في الشكل الجانبي (١,١٥)، حيث لدينا قيمة زاوية القطاع الدائري الأول (الموافق للعام 1423) والسادس (الموافق للعام 1428) تساويان:

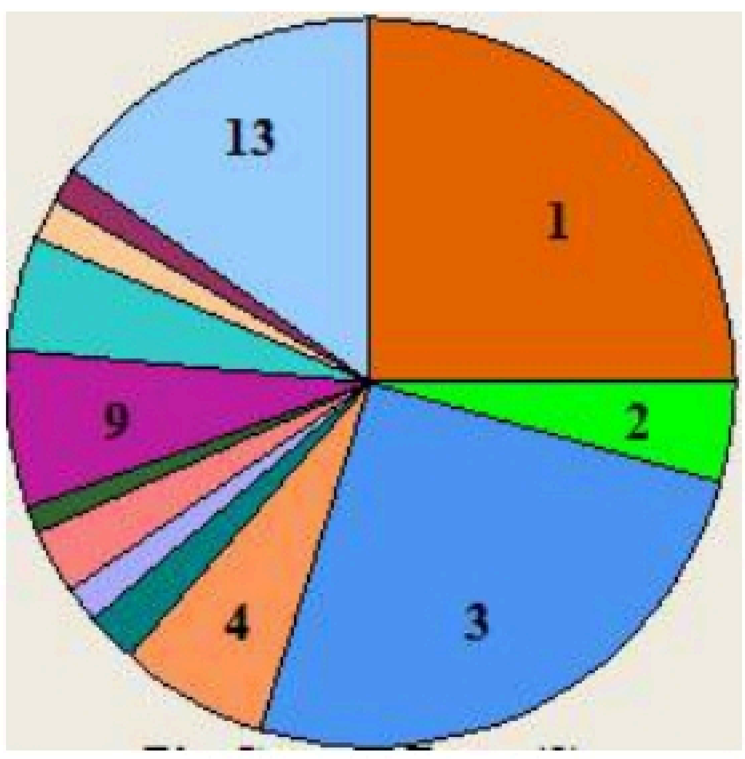


الشكل (١,١٥)

$$\alpha_1 = \frac{n_1}{n} \times 360 = \frac{1431012}{16276757} \times 360 = 31.65 = 31:39;00$$

$$\alpha_6 = \frac{n_6}{n} \times 360 = \frac{1707814}{16276757} \times 360 = 37.77 = 37:46;12$$

أخيراً إذا أردنا عرض البيانات (١,٣.ج) باستخدام تجزئة القرص (حيث لدينا حجم المجتمع $N = 27136977$)، فإنّنا نجد تمثيلها البياني في الشكل الآتي (١,١٦)، علماً أنّ قيمة الزاوية للقطاع الدائري الأول (الموافق لمنطقة الرياض) والسادس (الموافق لمنطقة الجوف) تساويان:




الشكل (١,١٦)

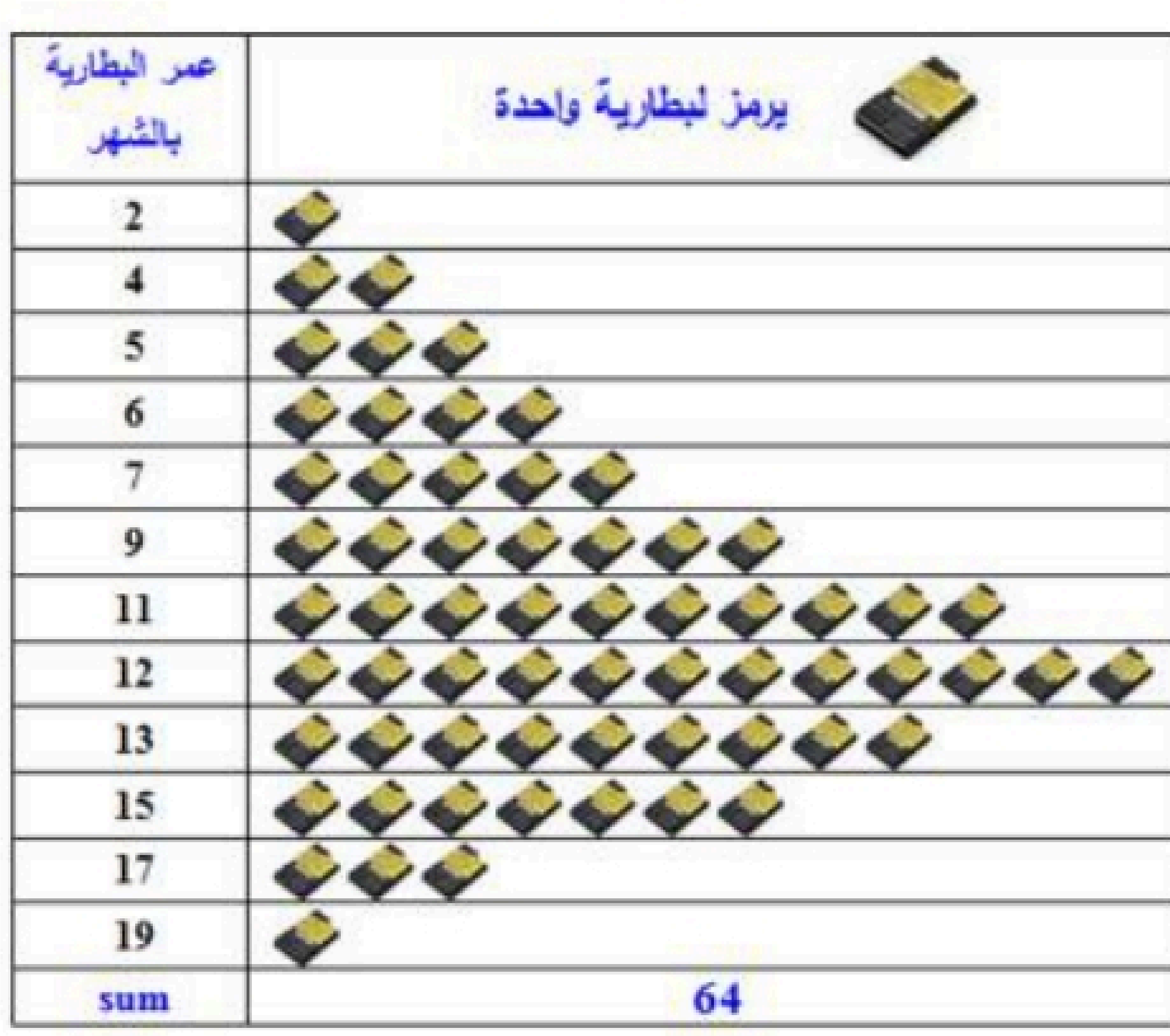
$$\alpha_1 = \frac{n_1}{N} \times 360 = \frac{6777146}{27136977} \times 360 = 89.91 = 89:54;36$$

$$\alpha_6 = \frac{n_6}{N} \times 360 = \frac{440009}{27136977} \times 360 = 5.84 = 5:50;24$$

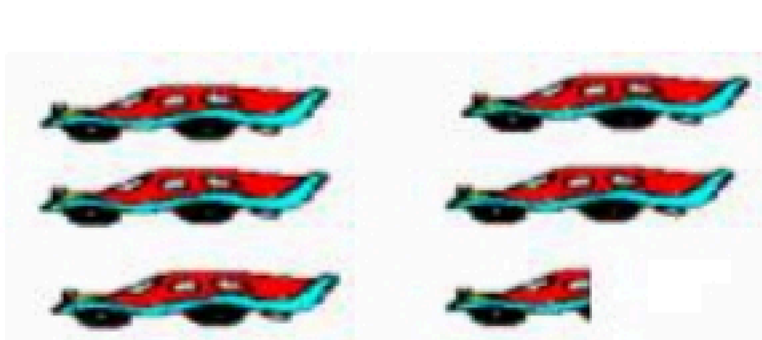
### (١,٣,٧) التمثيل بالصور (الطريقة التصويرية) Representation by Picture

في هذه الطريقة تستخدم صورة لتمثيل الشيء الخاضع للدراسة، ومن ثمَّ عرض هذه الصورة عدداً من المرّات يساوي التكرار للقيمة المقابلة، وهذه الطريقة شائعة الاستخدام في وسائط الإعلام كالصحف والمجلات، فعلى سبيل المثال لو أردنا عرض البيانات (١,٣ أ) وفقاً للطريقة التصورية فإنّنا سنجد لتمثيلها التصويري العرض الآتي:

الجدول (١,٦)




**مثال آخر**: لنفترض أنّه لدينا في مدينة ما 6500 سيّارة، وأردنا أنَّ نعبِّر عن ذلك بالطريقة التصويرية، فإنَّه برسم شكل سيّارة لكل 1000 سيارة فينتج لدينا العرض الجانبي لعدد السّيارات المُعطى.



## (١,٤) الجداول التكرارية ذوات الفئات

**Classes Tables**

في هذا النوع من الجداول يكون الاهتمام منصباً على عدد من الفئات (وهي عبارة عن فترات ذات طول موجب تماماً) بحيث يكون لكل منها سعة تساوي طول الفترة التي تمثّلها. إنَّ القيم المفردة للبيانات لا تكون موضع اهتمام في هذا النوع من الدراسة، ولذلك يُقال عن هذه البيانات إنَّها بيانات مُجمَّعة (أو بيانات مُجدولة) Grouped Data لأنَّه يتمّ تجميعها في جدول ذي فئات.

بناءً على ما ذُكر نلاحظ ضرورة تقديم بعض المفاهيم التي لا بدَّ منها لشرح الجداول التكرارية ذوات الفئات، وسنبدأها بالمفهوم الآتي

### (١ , ٤ , ١) تعريف (الفئة Class)

الفئة هي فترة يُصطلح عليها بحيث يكون لها طول موجب تماماً.

لقد ورد في المراجع تسميات عديدة للفئة منها: Interval أو Category أو Group أو Class. علماً أنَّ الحد الأدنى للفئة هي القيمة التي تمثّل الطرف الأيسر للفئة، وأمّا الطرف الأيمن للفئة فإنَّه القيمة التي تمثّل الحد الأعلى للفئة.

### (١ , ٤ , ٢) تعريف (سعة الفئة Length of Class)

سعة الفئة هو ذلك العدد الدّال على طول الفترة التي تشغلها الفئة وتساوي الفرق بين الحد الأعلى والأدنى للفئة.

سنرمز لسعة الفئة ذات الرقم $i$ بالرمز $C_i$، وفي حال تساوي السعة لجميع الفئات سنستخدم الرمز $C$ (بدون دليل للرمز) للدلالة على سعة الفئات.

### (١ , ٢ , ٤ , ١) ملاحظات

١ - إنَّ سعة الفئات تختار بحسب طبيعة المسألة التي هي قيد الدرس، وهنا يجب الانتباه إلى أنَّه من الممكن حصولنا على فئات خالية من البيانات، وهذا النوع من الفئات غير مرغوب فيه، ولذلك عند ظهور مثل هذه الحالة علينا زيادة أو إنقاص سعة الفئة، أو أن نقوم بدمجها مع فئة سابقة لها أو لاحقة بها بحيث نحصل على فئات غير خالية من البيانات. من جهة أخرى (وكما ذكرنا آنفًا) ليس بالضرورة أن تكون سعة جميع الفئات متساوية، ولكن في حال عدم تساوي سعة الفئات تصبح الدراسة أكثر صعوبةً وتعقيداً، ولهذا السبب فإنَّنا في دراستنا المقبلة لن نتعامل إلاَّ مع فئات متساوية السعة فقط، بمعنى أنَّ كل الجداول التكرارية التي سنأتي على ذكرها ستكون ذات فئات متساوية السعة ما لم ننوِّه بخلاف ذلك.

٢- يمكن التعديل على قيمة السعة للفئة (ولكنَّه ليس مُلزماً) بحيث نحصل على قيمة تجعل عمليّة تفريغ البيانات أكثر سهولة وكذلك العمليات الحسابية المتعلّقة بسعة الفئة أكثر يسراً، فعلى سبيل المثال لو كان لدينا $C = 2.37$ ، فعندئذٍ يُمكننا أخذ القيمة $C = 2.4$ ، $C = 2.3$ ، $C = 2.5$ أو أيّة قيمة أخرى قريبةً منها عوضاً عن تلك القيمة (ويفضّل التقريب بالزيادة وليس بالنقصان) وذلك حسب طبيعة الفارق بين الفئات ورأي الباحث في هذه القيم.

### (١ , ٤ , ٣) تعريف (تكرار الفئة Frequency of Class)

إنَّ تكرار الفئة ذات الرقم $i$ هو عدد سنرمز له بـ $f_i$ يدّل على عدد العناصر الموجودة في هذه الفئة (أو التي تنتمي إلى هذه الفئة).

الآن، ومن أجل بناء هذه الجداول يتمّ حساب أو اختيار قيم لسعات الفئات (سنأتي على كيفية تعيينها لاحقاً في هذا الفصل)، ومن ثمَّ يُعيّن في هذا الجدول عدّة أعمدة لكل منها مهمّة مُحدّدة، فمنها خاص بالفئات التي شكّلت، وآخر خاص بتعداد الفئات التي لدينا،

وثالث يُسجّل فيه عدد البيانات التي تنتمي لتلك الفئات المقابلة (التكرار)، وآخر لمهمّة أخرى وهكذا.... إنَّ هذا النوع من الجداول يستخدم (وكما ذكرنا قبل قليل) من أجل صبِّ (أو تجميع) البيانات الخام الكميّة عندما يكون حجم البيانات كبيراً، وعند بناء هذا النوع من الجداول لابدَّ من القيام بمجموعة إجراءات هي:

١ – التحقّق من أنَّ جميع البيانات من النوع الكميّ.

٢ – التحقّق من عدم وجود بيانات مفقودة.

٣– معرفة عدد الفئات المطلوبة أو سعاتها.

وهناك إجراءات أخرى سنأتي على ذكرها لاحقاً. لكن في الواقع يُميَّز بين نوعين من الجداول التكرارية ذوات الفئات وهما:

### (٤ , ٤ , ١) الجداول التكرارية ذوات الفئات المنفصلة

تكون الفئات في هذا النوع من الجداول منفصلة بعضها عن البعض الآخر بمقدار وحدة قياس أو جزء من وحدة قياس، وذلك حسب طبيعة البيانات والدقّة المستخدمة في أجهزة القياس التي أنتجت لنا البيانات، والعروض الجدولية الافتراضية الآتية توضّح لنا ذلك.

لنفترض أنّنا نقوم بوزن أشياء باستخدام ميزان دقّته 1 كغ، فعندئذ يمكننا تقديم البيانات الناتجة في جدول منفصل الفئات على النحو الآتي (قيم التكرارات في هذه الجداول هي قيم افتراضية):

الجدول (٧ , ١ .أ)

| رقم الفئة | الحدود للفئة | تكرار الفئة |
|---|---|---|
| 1 | 3 — 7 | 2 |
| 2 | 8 — 12 | 5 |
| 3 | 13 — 17 | 7 |
| 4 | 18 — 22 | 4 |
| 5 | 23 — 27 | 5 |
| Total | ------------- | 23 |

فنلاحظ هنا أنَّ الفاصل بين كل فئتين متتاليتين يساوي وحدة قياس كاملة.

أمّا إذا كنّا نقوم بوزن هذه الأشياء باستخدام ميزان دقّته 0.1 كغ، فعندئذ يمكننا أن نقدّم البيانات الناتجة في جدول منفصل الفئات على النحو الآتي:

الجدول (٧ , ١ .ب)

| رقم الفئة | حدود الفئة | تكرار الفئة |
|---|---|---|
| 1 | 3 — 7 | 2 |
| 2 | 7.1 — 11.1 | 4 |
| 3 | 11.2 — 15.2 | 8 |
| 4 | 15.3 — 19.3 | 6 |
| 5 | 19.4 — 23.4 | 3 |
| Total | ------------- | 23 |

حيث يُلاحظ هنا أنَّ الفاصل بين كل فئتين متتاليتين يساوي جزءًا عشريًّا واحدًا من وحدة القياس.

وبالمثل لو كنّا نقوم بوزن هذه الأشياء باستخدام ميزان دقّته 0.01 كغ، فعندئذ يمكننا أن نقدّم البيانات الناتجة في جدول منفصل الفئات على النحو الآتي:

الجدول (١,٧.ج)

| رقم الفئة | حدود الفئة | تكرار الفئة |
|---|---|---|
| 1 | 3 — 7 | 2 |
| 2 | 7.01 — 11.01 | 4 |
| 3 | 11.02 — 15.02 | 9 |
| 4 | 15.03 — 19.03 | 6 |
| 5 | 19.04 — 23.04 | 2 |
| Total | ---------------- | 23 |

وفي هذا الجدول يُلاحظ أنَّ الفاصل بين كل فئتين متتاليتين يساوي جزءًا مئويًا واحدًا من وحدة القياس.

**(١,٤,٤) ملاحظات:**

١- يُمكن جعل الفارق بين فئتين متتاليتين يساوي وحدة قياس إذا ما ضربت قيم حدود الفئات بمضاعف مناسب للعشرة، ومن ثمَّ يُمكننا النظر إلى الجداول ذوات الفئات المفصلة على أنّها من الشكل المقدَّم في الجدول الآتي (١,٨)، حيث يشير التكرار $f_1$، $f_2$، $f_3$ و ... في هذا النوع من الجداول إلى عدد القيم التي تنتمي إلى الفترة المغلقة $[a\ ,\ b]$، $[b\ +1,\ c]$، $[c\ +1,\ d]$ و... على الترتيب.

الجدول (١,٨)

| رقم الفئة | حدود الفئة | تكرار الفئة |
|---|---|---|
| 1 | $a$ — $b$ | $f_1$ |
| 2 | $b$+1 — $c$ | $f_2$ |
| 3 | $c$+1 — $d$ | $f_3$ |
| 4 | $d$+1 — $e$ | $f_4$ |
| 5 | $e$+1 — $f$ | $f_5$ |
| Total | ------------ | $\sum f_i$ |

٢- نشير هنا إلى أنَّ القيم التي تُصبّ في هذه الفئات هي تلك القيم التي نحصل عليها من قياس المشاهدات مباشرةً، ولذلك فهذه القيم ليست القيم الحقيقية للبيانات لأنَّها مجرَّدة من قيمة الخطأ المرتكب من أجهزة القياس. لذلك تُدعى هذه الفئات بـ **الفئات العمليّة** (أو التجريبيّة)، وحدودها يُطلق عليها اسم **الحدود العمليّة للفئات** Class Limits.

**(١,٤,٥) الجداول التكرارية ذوات الفئات المتّصلة**

تكون الفئات في هذا النوع من الجداول متصلة بعضها مع البعض الآخر بحيث تكون نهاية كل فئة هي بداية الفئة التالية، ومن ثمَّ يكون لهذا النوع من الجداول العرض الآتي:

جدول (١,٩)

| رقم الفئة | حدود الفئة | تكرار الفئة |
|---|---|---|
| 1 | $a \rightarrow b$ | $f_1$ |
| 2 | $b \rightarrow c$ | $f_2$ |
| 3 | $c \rightarrow d$ | $f_3$ |
| 4 | $d \rightarrow e$ | $f_4$ |
| 5 | $e \rightarrow f$ | $f_5$ |
| Total | ------------ | $\sum f_i$ |

وهنا نلاحظ أنَّ التكرار $f_1$ يكون دالّاً على عدد القيم التي تنتمي للفترة نصف المغلقة $[a , b)$، والتكرار $f_2$ دالّ على عدد القيم التي تنتمي للفترة نصف المغلقة $[b, c)$، وهكذا دواليك... ، ولذلك استخدمنا سهماً بدلاً من الشحطة للتأكيد على أنّه لدى صبِّ البيانات الخام في هذا النوع من الجداول، فإنّه إذا وافقت إحدى قيم البيانات للحد الأعلى للفئة فإنّها تُصبّ (أو تُفرَّغ) في الفئة التالية.

### (١,٥,٤,١) ملاحظات

١- يُلجأ عادة إلى هذا النوع من الجداول إذا كانت دقّة أجهزة القياس المستخدمة في توليد البيانات عاليةً جداً لدرجة يمكن معها إهمال الخطأ المرتكب من هذه الأجهزة، أو أن نكون غير مهتمين بقيمة الخطأ المرتكب لدى توليد البيانات حيث يُنظر إلى القيم التي تمّ الحصول عليها من قياس المشاهدات على أنّها القيم الحقيقيّة للبيانات، ولذلك تُدعى هذه الفئات بـ **الفئات الفعليّة** (أو الحقيقيّة)، وحدود هذه الفئات تُدعى **الحدود الفعليّة للفئات** Class Boundaries.

٢- نشير إلى أنّه إذا وجدت قيم تساوي قيمة الحد الأعلى للفئة الأخيرة في الجدول التكراري فإنّنا نقوم بضمّها للفئة الأخيرة، وأمّا إذا وجدت قيم تساوي الحد الأعلى للفئة الأخيرة وأخرى أكبر من الحد الأعلى للفئة الأخيرة فإنّنا نقوم بإضافة فئة جديدة تلي آخر فئة لتشتمل على ما تبقى من بيانات بما فيها القيم المساوية للحد الأعلى للفئة الأخيرة.

٣- إذا ظهرت لدينا فئات خالية من البيانات فإنّنا نجري تعديلاً على عدد الفئات (ومن ثمّ على سعة الفئات) بحيث ننفي وجود أي فئة خالية من البيانات في جدول التوزيع التكراري.

### (١,٤,٦) حساب سعات الفئات وعددها من أجل جدول تكراري ذو فئات

بفرض أنّه لدينا مجموعة بيانات خام $x_1$، $x_2$، ... و $x_n$ ، فعندئذ لحساب عدد الفئات اللازمة لبناء جدول تكراري ذي فئات خاص بهذه البيانات نتبع الخطوات الآتية:

١- حساب **المدى** Range لهذه البيانات، وسنرمز له بـ $\mathcal{R}$ ، ويُعرَّف بالعلاقة الآتية:

$$\mathcal{R} = x_\ell - x_S \qquad [1,3]$$

علماً أنَّ $x_\ell$ و $x_S$ ترمزان لأكبر وأصغر قيمة في البيانات على الترتيب.

٢- إذا فُرِض علينا عدد مُعيَّن من الفئات وليكن $k$ مثلاً، فعندئذ تُحسب **السعة** $\mathbf{C}$ لكل فئة من الجدول بوساطة العلاقة الآتية:

$$\mathbf{C} = \frac{\mathcal{R}}{k} \qquad [1,4]$$

٣- إذا قُدِّمت لنا السعة $\mathbf{C}$ الواجب استخدامها لكلّ فئة في الجدول، فعندئذ يُحسب عدد الفئات $k$ اللازمة لبناء هذا الجدول بوساطة العلاقة الآتية:

$$k = \left\lceil \frac{\mathcal{R}}{\mathbf{C}} \right\rceil \qquad [1,5]$$

علماً أنَّ $\lceil x \rceil$ هو أصغر عدد صحيح أكبر من أو يساوي $x$.

٤ – إذا لم نُعط سعة الفئات، ولم يُفرض علينا عدد معيّن من الفئات، فعندئذ تُستخدم العلاقة الآتية لتعيين عدد الفئات:

$$k = \lfloor 3.322 \log n \rfloor \qquad [1,6]$$

وسوف ندعوه **العدد الافتراضي** Default Number للفئات، علماً أنَّ $n$ هو عدد البيانات الخام التي ينبغي صبّها في الجدول و $\lfloor x \rfloor$ هو أكبر عدد صحيح أصغر من أو يساوي $x$. بعد ذلك تُحسب سعة الفئة من خلال العلاقة [1,4] مع أخذ $k = \lfloor 3.322 \log n \rfloor$، حيث سندعو هذه السعة بـ **السعة الافتراضية** Default Class width.

نُشير هنا إلى أنَّ العلاقة [1,6] تُستخدم في بعض البرامج الإحصائية لتعيين العدد الافتراضي للفئات ومنها برنامج Easy Fit، حيث يُلاحظ هنا التناسب الطردي لعدد الفئات مع عدد البيانات $n$ (أو مجموع التكرارات).

**أمثلة (١, ٦, ٤, ١)**

١ -لنقم بصبّ مجموعة البيانات (٣, ١.أ) في جدول تكراري.

من أجل ذلك يجب علينا حساب المدى $\mathcal{R}$ لتلك البيانات، فنجد أنَّ $\mathcal{R}=19-2=17$، وبما أنَّه لم يُقدَّم لنا عدد الفئات المطلوب استخدامها فإنَّنا سنستخدم العدد الافتراضي للفئات فنجده يساوي:

$$k = \lfloor 3.322 \log n \rfloor = \lfloor 3.322 \log 64 \rfloor = \lfloor 6.00 \rfloor = 6$$

ومن ثَّم تكون السعة الافتراضية للفئات التي سنستخدمها هي:

$$\mathbf{C} = \frac{\mathcal{R}}{k} = \frac{17}{6} = 2.83$$

وعلى سبيل التبسيط سنأخذ $\mathbf{C} = 2.9$ (كان من الممكن أن نأخذ $\mathbf{C} = 3$ أيضاً)، والفاصل بين كل فئتين منفصلتين متتاليتين يساوي 0.1 (أي افترضنا أنَّ دقَّة القياس **0.1**). بالطبع يجب الانتباه هنا إلى القيمة التي سنأخذها كفاصل بين كل فئتين متتاليتين، فلو أخذنا (على سبيل المثال) هذه القيمة مساويةً لوحدة قياس صحيحة (أي بمقدار **1**) فإنَّ الفئة السادسة ستصبح خارج نطاق البيانات (أي ستصبح فئة خالية).

الآن بصبّ بيانات المجموعة (٣, ١.أ) في جدول تكراري وفقاً لما شرحناه آنفاً سنحصل على الجدول التكراري (١٠, ١.أ) في حال استخدامنا فئات منفصلة، وأمّا في حال استخدامنا فئات متصلة فإنَّه سيكون للبيانات (٣, ١.أ) الجدول التكراري (١٠, ١.ب).

الجدول (١٠, ١.أ)

| رقم الفئة | حدود الفئة | تكرار الفئة |
|---|---|---|
| 1 | 2 — 4.9 | 3 |
| 2 | 5 — 7.9 | 12 |
| 3 | 8 — 10.9 | 7 |
| 4 | 11 — 13.9 | 31 |
| 5 | 14 — 16.9 | 7 |
| 6 | 17 — 19.9 | 4 |
| **Total** | ------------ | **64** |

الجدول (١٠ ,١.ب)

| رقم الفئة | حدود الفئة | تكرار الفئة |
|---|---|---|
| 1 | $2 \rightarrow 4.9$ | 3 |
| 2 | $4.9 \rightarrow 7.8$ | 12 |
| 3 | $7.8 \rightarrow 10.7$ | 7 |
| 4 | $10.7 \rightarrow 13.6$ | 31 |
| 5 | $13.6 \rightarrow 16.5$ | 7 |
| 6 | $16.5 \rightarrow 19.4$ | 4 |
| Total | ------------ | 64 |

فلاحظ هنا أنَّه من أجل هذا المثال لم تتغيّر قيم التكرارات في كلا الحالتين، ولكن هذا ليس قاعدة عامّة، فمن الممكن أن تتغيّر قيم التكرارات في حال الانتقال من جدول تكراري منفصل الفئات إلى جدول تكراري متصل الفئات والمثال الآتي يوضِّح لنا ذلك.

٢- لنأخذ البيانات الخام الآتية:

| | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | 2 | 11 | 11 | 2 | 9 | 9 | 2 | 9 | 6 | 5 | 5 | 5 | 11 | 15 | 17 | 15 | 13 |
| 6 | 9 | 5 | 7 | 7 | 4 | 4 | 4 | 5 | 5 | 7 | 4 | 13 | 4 | 13 | 19 | 13 | 19 |

ولنقم بصبّ هذه البيانات في جدول تكراري بفئات منفصلة ذات سعة تساوي 3 فإنَّه سيكون للبيانات المُعطاة الجدول التكراري (١١ ,١.أ)، وأمّا إذا صُبَّت البيانات في جدول تكراري بفئات متصلة ذات سعة تساوي 3 فإنَّنا سنحصل على الجدول التكراري (١١ ,١ ب) للبيانات المُعطاة.

الجدول (١١ ,١.أ)

| رقم الفئة | حدود الفئة | تكرار الفئة |
|---|---|---|
| 1 | 2 — 5 | 14 |
| 2 | 6 — 9 | 10 |
| 3 | 10 — 13 | 7 |
| 4 | 14 — 17 | 3 |
| 5 | 18 — 21 | 2 |
| Total | ------------ | 36 |

الجدول (١١ ,١.ب)

| رقم الفئة | حدود الفئة | تكرار الفئة |
|---|---|---|
| 1 | $2 \rightarrow 5$ | 8 |
| 2 | $5 \rightarrow 8$ | 11 |
| 3 | $8 \rightarrow 11$ | 5 |
| 4 | $11 \rightarrow 14$ | 7 |
| 5 | $14 \rightarrow 17$ | 2 |
| 6 | $17 \rightarrow 20$ | 3 |
| Total | ------------ | 36 |

حيث يتّضح لنا تغيّر قيم التكرارات وعدد الفئات أيضاً.

### (١,٤,٧) جداول التوزيع التكرارية Frequency Distribution Tables:

في الواقع يمكن تمديد الجداول التكرارية ذوات الفئات بحيث تصبح أكثر فاعليّة في تنفيذ بعض العمليّات الحسابيّة الخاصّة بالبيانات وكذلك توضيح شكل التوزيع للبيانات، ومن ثمَّ إضفاء صفة العموميّة على هذه الجداول. لذلك فإنَّ كل ما تمّ شرحه سابقاً بشأن الجداول التكراريّة يبقى ساري المفعول من أجل هذه الجداول باستثناء ما يشار إليه في الدراسة التالية. إنَّ هذه الجداول تُعرف باسم **جداول التّوزيع التّكراريّة** ولها نماذج مختلفة ولكنَّ معظم هذه الجداول تحتوي على أعمدة من أجل:

**١- الرقم المتسلسل للفئات Number of Classes:**

في هذا العمود يدوّن فيه أرقام متسلسلة دالّة على رقم الفئات في الجدول (كما سبق واستخدم في الجداول التكراريّة)، وينتهي الترقيم مع القيمة التي تُحدِّد عدد الفئات حيث كنّا قد ذكرنا سابقاً كيفية تعيين عدد الفئات لجدول توزيع تكراري، ولكن من أجل هذا النوع من الجداول التكرارية يُراعى عند اختيار عدد الفئات ما يلي:

أ- يجب ألّا يقلّ عدد الفئات عن خمس، وهذا يوافق الحالة التي يكون فيها عدد البيانات المُراد صبّها لا يقل عن 32 قيمة وفقاً للعدد الافتراضي للفئات $k = \lfloor 3.322\log n \rfloor$ حيث لدينا $n = 10^{\frac{5}{3.322}} \approx 32$، وهذا يتضّح لنا من خلال العلاقتين الآتيتين أيضاً:

$$k = \lfloor 3.322 \ \log 31 \rfloor = \lfloor 4.954304 \rfloor = 4$$

$$k = \lfloor 3.322 \ \log 32 \rfloor = \lfloor 5.000108 \rfloor = 5$$

ولذلك سنقول عن بيانات مُعطاة إنّها قليلة العدد (أو عيِّنة صغيرة الحجم) إذا كان $n < 32$ ، وخلاف ذلك نعدّها كثيرة العدد أو عيِّنة كبيرة الحجم.

ب- يجب ألّا يزيد عدد الفئات عن العشرين، وهذا يوافق الحالة التي يكون فيها عدد البيانات الخام المراد صبّها لا يزيد عن $10^6$ وفقاً للعدد الافتراضي للفئات حيث لدينا $n = 10^{\frac{20}{3.322}} \approx 10^6$ ، وسبب ضرورة تحقّق البندين السّابقين هو أنّه في الحالة الأولى يصبح الجدول غير عمليّ في الدراسة الإحصائية للظاهرة التي هي قيد البحث وخاصة إذا كان حجم العيِّنة كبيراً بما فيه الكفاية، وفي الحالة الثانية يصبح الجدول مجهداً في دراسته، ولكنّه لبعض الضرورات يمكن تجاهل البندين السّابقين مثل الحالات التي يكون فيها حجم البيانات التي نودّ صبّها كبير جداً (كالتعداد السكاني الشامل)، أو إذا كنّا بحاجة لصبّ مجموعة بيانات بحجم 30 أو 31 في جدول توزيع تكراري بخمس من الفئات فقط.

**٢- الحدود العمليّة (أو التجريبية) للفئات Class Limits:**

يُدوّن في هذا العمود الفئات المتتالية بشكلها المنفصل (فئات منفصلة) على النحو الذي تمَّ شرحه سابقاً في فقرة الجداول التكرارية ذوات الفئات المنفصلة، وموضّحاً فيها قيم حدَّي كل فئة، وتؤخذ عادة أصغر قيمة للبيانات الخام كحد أدنى للفئة الأولى.

**٣- الحدود الفعليَّة (أو الحقيقية) للفئات Class Boundaries:**

يُدوّن في هذا العمود فئات تكون فيها قيم حدّها الأدنى أقلّ بمقدار نصف وحدة عن حد الفئة العمليّة (حسب الوحدة المستخدمة في دقّة القياس كفاصل بين كل فئتين، حيث قمنا بتوضيح ذلك سابقاً)، وأمّا حدّها الأعلى فإنّه يزيد بمقدار نصف وحدة عن حد الفئة العمليّة، وعلى ألاَّ تُأخذ قيمة الحدّ الأعلى للفئة كجزء من الفئة الفعليّة، أي إنَّ الفئات النّاتجة عن هذه العمليّة ستكون على شكل فترات نصف مفتوحة من الشكل $[a , b)$. لذلك تصبح الفئات المتتالية في هذا العمود متصلة بعضها مع البعض الآخر. إنَّ السبب في أخذ هذا النوع من الفئات هو أنَّه لدى عمليّة تحصيل البيانات سيكون لدينا خطأً مُرتكب من وسائل القياس المُعتمدة (مهما بلغت من

دقة)، ولذلك اصطلح على أنَّ نصف وحدة قياس ستغطي الخطأ المرتكب زيادةً أو نقصاناً، ولذلك ستحتوي هذه الفئات على القيم الحقيقية (أوالفعليّة) للبيانات ، ولهذا السبب دُعيت هذه الفئات بـ **الفئات الفعليَّة**. هنا نشير إلى أنَّ سعة الفئة الفعليّة **C** تحسب من أجل هذا النوع من الجداول باستخدام العلاقة الآتَية عندما يكون عدد الفئات المعطاة يساوي $k$:

$$\mathbf{C} = \frac{\mathcal{R} + one\ \ measuring\ \ unit}{k}$$

و أمّا في حال كان عدد الفئات غير مُعطى فإنَّنا نستخدم العلاقة الآتية لحساب السعة:

$$\mathbf{C} = \frac{\mathcal{R} + one\ \ measuring\ \ unit}{\lfloor 3.322 \log n \rfloor}$$

إنَّ إضافة وحدة القياس إلى قيمة المدى لأجل تغطية الخطأ الأعظمي المرافق لقياس أكبر وأصغر قيمة في البيانات.

بعد ذلك نقوم بتعيين حدود الفئات الفعلّية كما يلي:

أ- نجعل الحد الأدنى لأول فئة فعلّية مساوياً للقيمة الناتجة عن طرح نصف وحدة (حسب الوحدة المستخدمة في دقّة القياس) من أصغر قيمة في البيانات.

ب- نضيف قيمة **C** إلى الحد الأدنى للفئة الأولى فنحصل على الحد الأعلى للفئة الفعلّية الأولى.

ج- نجعل الحد الأدنى للفئة الفعلّية التالية (الثانية ) مساوياً للحد الأعلى للفئة الفعلّية السابقة (الأولى)، ومن ثمَّ نضيف قيمة **C** إلى الحد الأدنى لهذه الفئة فنحصل على الحد الأعلى للفئة الفعلّية الثانية.

د- نقوم بتطبيق الفقرة السابقة (ج) من أجل جميع الفئات المتبقّية فنحصل على الحدود الدنيا والعليا للفئات الفعلّية لجدول التوزيع التكراري.

نشير هنا إلى أنَّه في كثير من الحالات يُستغنى عن عرض عمود حدود الفئات العمليّة Class Limit لأنَّ صبَّ (أو تفريغ) البيانات الخام في هذا النوع من الجداول يتمّ في الفئات الفعليَّة حصراً، وكذلك الدراسات اللاحقة لسلوك بيانات جداول التوزيع التكرارية من عروض بيانية ومقاييس عددية تعتمد جميعها على عمود الفئات الفعليَّة فقط.

**٤- مراكز الفئات Midpoints:**

يُدوّن في هذا العمود، ومقابل الفئة التي هي قيد المسح، قيمة تدلَّ على مركز هذه الفئة، وقيمته تساوي نصف مجموع حدّيها الأعلى والأدنى، علماً أنَّه يُنظر إلى قيمة مركز الفئة على أنّه القيمة الممثِّلة للفئة، وسنرمز بـــ $x_i$ لقيمة مركز الفئة ذات الرقم $i$ .

**٥- تعداد عناصر الفئات Tally:**

يرسم في هذا العمود، ومقابل الفئة التي هي قيد المسح، خطًّا من الشكل | عن كل بيان ينتمي لهذه الفئة التي هي قيد المسح، وبحيث تُجمَّع كل خمسة خطوط على شكل حزمة ~~||||~~، وغالباً ما يُحْذف هذا العمود لأنَّ عمود تكرار الفئات (الذي سيقدّم في الفقرة التالية) يُغْني عن هذا العمود.

**٦- تكرار الفئات Frequency:**

في هذا العمود يُدوّن التكرار للفئة التي هي قيد المسح على النحو الذي ذكرناه سابقاً، وسنرمز بـــ $f_i$ لقيمة تكرار الفئة ذات الرقم $i$ .

**٧- التّكرار النسبي للفئات Relative Frequency:**

في هذا العمود يُدوّن التكرار النسبي للفئة التي قيد المسح، وقيمته هي حاصل قسمة تكرار الفئة المعنية على المجموع الكلي للتكرارات، وغالباً يُحْذف هذا العمود من جدول التوزيع التكراري لأنَّ قيمه تُستنتج من القيم الموجودة في عمود تكرار الفئات.

**٨- التّكرار المئوي (تكرار النسبة المئوية) للفئات Percentage Frequency:**

في هذا العمود يُدوّن التكرار المئوي (أو تكرار النسبة المئوية) للفئة التي هي قيد المسح، وقيمة هذا التكرار ينتج من خلال ضرب التكرار النسبي للفئة بالعدد 100 ويقرأ كنسبة مئوية، وغالباً يُحْذف هذا العمود؛ لأنَّه يستنتج من القيم الموجودة في عمود تكرار الفئات أيضاً.

**٩- التّكرار المتجمِّع الصّاعد للفئات Ascending Cumulative Frequency:**

في هذا العمود، ومقابل الفئة التي هي قيد المسح، يُدوّن عدد البيانات التي قيمها أقل من قيمة الحد الأعلى لهذه الفئة، ومن ثَمَّ يكون التكرار المتجمِّع الصاعد لفئة يساوي مجموع تكرار الفئة مع جميع التكرارات للفئات السابقة لها، وسنرمز بـ $F_i$ لقيمة التكرار المتجمِّع الصّاعد للفئة ذات الرقم $i$.

**١٠- التّكرار المتجمِّع الهابط للفئات Descending Cumulative Frequency:**

في هذا العمود، ومقابل الفئة التي هي قيد المسح، يُدوّن عدد البيانات التي قيمها أكبر من قيمة الحد الأدنى لهذه الفئة، وهو نادر الاستخدام، ولذلك لايدرج في جدول التوزيع التكراري إلاّ عند الحاجة له.

بهذا يكون لجدول التوزيع التكراري العرض الآتي:

الجدول (١٢ , ١)

| التكرار المتجمِّع الهابط للفئة | التكرار المتجمِّع الصاعد للفئة $F_i$ | التكرار المئوي للفئة | التكرار النسبي للفئة | تكرار الفئة $f_i$ | تعداد الفئة | مركز الفئة $x_i$ | الحدود الفعليّة للفئة | الحدود العمليّة للفئة | رقم الفئة $i$ |
|---|---|---|---|---|---|---|---|---|---|
| … | … | … | … | … | … | … | … | … | **1** |
| … | … | … | … | … | … | … | … | … | **2** |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| المجموع | --------- | | | | ------ | ------ | ------ | ------- | **Total** |

**(١ , ٧ , ٤ , ١) ملاحظات**

١- نوكِّد هنا أنَّه عند صبّ (أو تفريغ ) البيانات الخام في جدول توزيع تكراري فإنَّه يتمّ بناءً على قيم الحدود الفعليَّة للفئات، بمعنى أنَّ التفريغ يتمّ في الفئات الفعليَّة وليس في الفئات العمليّة، وتُحسب سعة الفئة الفعليَّة C باستخدام العلاقة [1,4] مع أخذ قيمة عدد الفئات $k = \lfloor 3.322 \log n \rfloor$ .

٢-إذا طُلِب منّا صبّ البيانات الخام في جدول توزيع تكراري بفئات عمليّة ذات سعة مُحدَّدة فإنَّنا نقوم بتعيين الفئات العمليّة في الجدول، ومن ثَمَّ نستنتج منها الفئات الفعليّة التي ستصبّ فيها البيانات.

٣- عندما تكون أصغر الفروق بين البيانات أعدادًا صحيحة فإنَّه يمكن أخذ قيمة الفاصل بين الفئة العمليّة والفئة التالية لها من الشكل $10^k$ مع $k = 0,1,2,...,n$ و $n$ عدد صحيح غير سالب، فعلى سبيل المثال إذا كانت أصغر الفروق بين البيانات من مرتبة الآحاد فإنَّنا نستخدم الفاصل بين الفئة العمليّة والفئة التالية لها بمقدار يساوي الواحد (أي $10^0$) وتكون الدقّة المعتمدة في هذه الحالة تساوي $\pm 0.5$ ، وأمّا إذا كانت أصغر الفروق بين البيانات من مرتبة العشرات فإنَّنا نستخدم الفاصل بين الفئة العمليّة والفئة التالية لها بمقدار يساوي عشرة (أي $10^1$) وتكون الدقّة المعتمدة في هذه الحالة تساوي $\pm 5$ ، وهكذا دواليك.... أمّا إذا كانت هناك بعض القيم غير الصحيحة فعندئذ يجب النظر في القيمة التي ستستخدم كفاصل بين الفئات وفقاً لما تسمح به المسألة المدروسة.

٥- إضافة لما سبق فإنَّه من الممكن أن يتضمن جدول التوزيع التكراري السّابق بعض الأعمدة الأخرى مثل التكرار المتجمِّع الصاعد النسبيّ والتكرار المتجمِّع الهابط النسبيّ، ولكن هذه الأعمدة تُضاف إلى جدول التوزيع التكراري عند الحاجة لها فقط.

### (١ , ٤ , ٧ , ٢) أمثلة

١- لتكن لدينا مجموعة البيانات الخام الآتية:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 7 | 5 | 2 | 9 | 7 | 11 | 7 | 4 | 3 |
| 3 | 10 | 3 | 5 | 10 | 4 | 6 | 8 | 4 | 8 |
| 6 | 9 | 6 | 5 | 6 | 8 | 5 | 6 | 5 | 9 |
| 7 | 9 | 7 | 9 | 7 | 7 | 11 | 7 | 9 | 10 |

ولنقم بصبِّ هذه البيانات في جدول توزيع تكراري.

**التنفيذ:** من أجل ذلك لنقم أولاً بحساب المدى لهذه البيانات، حيث لدينا من العلاقة [1,3] ما يلي:

$$\mathcal{R} = x_\ell - x_s = 11 - 2 = 9$$

وبما أنّه لم نُعطَ سعة الفئات، ولم يُفرض علينا عدد معيّن من الفئات، فعندئذ تُستخدم العلاقة [1,6] لحساب العدد الافتراضي للفئات، فيكون لدينا:

$$k = \lfloor 3.322 \log 40 \rfloor = \lfloor 5.22 \rfloor = 5$$

ومن ثمَّ نُحسب السعة الافتراضيّة C لكلّ فئة فعليّة من الجدول بوساطة العلاقة [1,4] حيث لدينا:

$$\mathbf{C} = \frac{\mathcal{R} + one\ measuring\ unit}{k} = \frac{9+1}{5} = 2$$

وهكذا يكون الحد الأدنى لأول فئة فعليّة يساوي $2 - 0.5 = 1.5$ ، ولكن وفقاً لهذه السعة من الممكن أن تبقى لدينا قيم خارج نطاق الفئة الفعليّة الأخيرة، ولذلك إمّا أن نزيد من سعة الفئة قليلاً أو نضطر إلى إضافة فئة سادسة، وهذا يوضِّح لنا الجدول الآتي:

الجدول (١٣ , ١ .أ)

| من أجل C = 1.8 | | من أجل C = 2 | |
|---|---|---|---|
| الحد الأدنى للفئة | الحد الأعلى للفئة | الحد الأدنى للفئة | الحد الأعلى للفئة |
| 1.5 | 3.3 | 1.5 | 3.5 |
| 3.3 | 5.1 | 3.5 | 5.5 |
| 5.1 | 6.9 | 5.5 | 7.5 |
| 6.9 | 8.7 | 7.5 | 9.5 |
| 8.7 | 10.5 | 9.5 | 11.5 |

وباتباع الخطوات التي قمنا بشرحها سابقاً فإنَّنا سنميِّز التفريغ للبيانات باستخدام لون خاص لكلّ فئة (وذلك على سبيل التوضيح فقط) فيكون لدينا العرض الآتي للبيانات المُعدّة للتفريغ:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 7 | 5 | 2 | 9 | 7 | 11 | 7 | 4 | 3 |
| 3 | 10 | 3 | 5 | 10 | 4 | 6 | 8 | 4 | 8 |
| 6 | 9 | 6 | 5 | 6 | 8 | 5 | 6 | 5 | 9 |
| 7 | 9 | 7 | 9 | 7 | 7 | 11 | 7 | 9 | 10 |

وبتفريغ هذه البيانات سنحصل على جدول التوزيع التكراري الآتي (١٣ , ١.ب) الذي يحتوي على الفئات الفعليّة، و الفئات العمليّة، ومراكز الفئات، والتعدادات، والتكرارات بأنواعها المشروحة سابقاً (لقد تمّ تجزئة الجدول إلى جدولين بسبب كبر حجم الجدول وضيق الحيِّز المتاح له).

الجدول (١٣ , ١.ب)

| رقم الفئة | لحدود العمليّة للفئة | لحدود الفعليّة للفئة | مركز الفئة | تعداد الفئة | تكرار الفئة | التكرار النسبي للفئة |
|---|---|---|---|---|---|---|
| 1 | 2 – 3 | 1.5 → 3.5 | 2.5 | 𝍸 | 5 | 5/40 = 0.125 |
| 2 | 4 – 5 | 3.5 → 5.5 | 4.5 | 𝍸 ||| | 8 | 8/40 = 0.200 |
| 3 | 6 – 7 | 5.5 → 7.5 | 6.5 | 𝍸 𝍸 ||| | 13 | 13/40 = 0.325 |
| 4 | 8 – 9 | 7.5 → 9.5 | 8.5 | 𝍸 |||| | 9 | 9/40 = 0.225 |
| 5 | 10 – 11 | 9.5 → 11.5 | 10.5 | 𝍸 | 5 | 5/40 = 0.125 |
| Total | ---------- | ----------- | ----- | ------ | 40 | 1 |

الجدول (١٣ , ١.ب)

| رقم الفئة | لحدود الفعليّة للفئة | تكرار الفئة | التكرار المئوي للفئة % | التكرار المتجمّع الصاعد الفئة | التكرار المتجمّع الهابط الفئة |
|---|---|---|---|---|---|
| 1 | 1.5 → 3.5 | 5 | 0.125×100 | 5 | 5+8+13+9+5=40 |
| 2 | 3.5 → 5.5 | 8 | 0.200×100 | 5+8=13 | 8+13+9+5=35 |
| 3 | 5.5 → 7.5 | 13 | 0.325×100 | 5+8+13=26 | 13+9+5=27 |
| 4 | 7.5 → 9.5 | 9 | 0.225×100 | 5+8+13+9=35 | 9+5=14 |
| 5 | 9.5 → 11.5 | 5 | 0.125×100 | 5+8+13+9+5=40 | 5 |
| Total | ------------ | 40 | 100 % | المجموع | المجموع |

الجدولين السابقين هما بمثابة جدول واحد

٢- صبّ مجموعة البيانات (٣ , ١.أ) في جدول توزيع تكراري بحيث تكون سعة الفئة العمليّة فيه تساوي $C = 2.5$، وبحيث يكون الفاصل بين الفئات العمليّة يساوي الواحد تماماً.

**التنفيذ:** نلاحظ هنا أنَّه فرض علينا سعة الفئات العمليّة والقيمة التي تفصل بينها، ولذلك علينا أولاً أن نعيّن الفئات العمليّة ومن ثمَّ نستنتج منها الفئات الفعليّة.

من أجل ذلك سنأخذ أصغر قيمة في مجموعة البيانات (٣ , ١.أ) كحد أدنى للفئة العمليّة الأولى فتكون الفئة العمليّة الأولى ممثّلة بالفترة [4.5 , 2] والثانية بالفترة [8 , 5.5] وهكذا دواليك...، وبما أنَّ الدقّة المعتمدة لدى الحصول على البيانات تساوي ±0.5، فإنَّه سيكون لتلك البيانات جدول التوزيع التكراري المقدَّم في الجدول الآتي (١٣ , ١.ج) (الجدولين هما بمثابة جدول واحد، وجزِّئا لضيق الحيِّز المخصّص لعرضهما في جدول واحد).

الجدول (١٣ , ١ .ج)

| رقم الفئة | الحدود العمليّة للفئة | الحدود الفعليّة للفئة | مركز الفئة | تعداد الفئة | تكرار الفئة | التكرار النسبي للفئة |
|---|---|---|---|---|---|---|
| 1 | 2.0 – 4.5 | 1.5 → 5 | 3.25 | \|\|\| | 3 | 3/64 = 0.0469 |
| 2 | 5.5 – 8.0 | 5 → 8.5 | 6.75 | ~~\|\|\|\|~~ ~~\|\|\|\|~~ \|\| | 12 | 12/64 = 0.1875 |
| 3 | 9.0 – 11.5 | 8.5 → 12 | 10.25 | ~~\|\|\|\|~~ ~~\|\|\|\|~~ ~~\|\|\|\|~~ \|\| | 17 | 17/64 = 0.2656 |
| 4 | 12.5 – 15 | 12 → 15.5 | 13.75 | ~~\|\|\|\|~~ ~~\|\|\|\|~~ ~~\|\|\|\|~~ ~~\|\|\|\|~~ ~~\|\|\|\|~~ \|\|\| | 28 | 28/64 = 0.4375 |
| 5 | 16 – 18.5 | 15.5 → 19 | 17.25 | \|\|\| | 3 | 3/64 = 0.0469 |
| 6 | 19.5 – 22 | 19 → 22.5 | 20.75 | \| | 1 | 1/64 = 0.0156 |
| Total | -------------- | ------------ | ----- | --------------------------- | 64 | 1 |

الجدول (١٣ , ١ .ج).

| رقم الفئة | الحدود الفعليّة للفئة | تكرار الفئة | التكرار المئوي للفئة % | التكرار المتجمّع الصاعد الفئة | التكرار المتجمّع الهابط الفئة |
|---|---|---|---|---|---|
| 1 | 1.5 → 5 | 3 | 0.0469×100 | 3 | 3+12+17+28+3+1=64 |
| 2 | 5 → 8.5 | 12 | 0.1875×100 | 3+12=15 | 12+17+28+3+1=61 |
| 3 | 8.5 → 12 | 17 | 0.2656×100 | 3+12+17=32 | 17+28+3+1=49 |
| 4 | 12 → 15.5 | 28 | 0.4375×100 | 3+12+17+28=60 | 28+3+1=32 |
| 5 | 15.5 → 19 | 3 | 0.0469×100 | 3+12+17+28+3=63 | 3+1=4 |
| 6 | 19 → 22.5 | 1 | 0.0156×100 | 3+12+17+28+3+1=64 | 1 |
| Total | ------------- | 64 | 100 % | المجموع | المجموع |

الجدولين السابقين هما بمثابة جدول واحد

### (٣ , ٧ , ٤ , ١) ملاحظات:

١ – بناءً على التنويهات التي وردت سابقاً يمكن اختزال جدول التوزيع التكراري (١٢ , ١) بحيث يحتوي الفئات الضروريّة منها فقط، ومن ثَمَّ استخدام عدد أقلّ من الأعمدة ( **وهذا بدوره يجعل الدراسة أكثر يُسراً وبساطةً**)، حيث يصبح لدينا جدول توزيع تكراري مُخْتَزل له العرض الآتي:

الجدول (١٤ , ١)

| رقم الفئة Number of Classes | الحدود الفعليّة للفئة Class Boundaries | مركز الفئة Midpoints | تكرار الفئة Frequency | التكرار المتجمّع الصاعد للفئة Ascending Cumulative Frequency |
|---|---|---|---|---|
| 1 | … | … | … | … |
| 2 | … | … | … | … |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| Total | -------------- | -------- | | ------------------ |

علماً أنَّ جدول التوزيع التكراري (١٢ , ١) المقدّم سابقاً يُنظر إليه كجدول نموذجي لعرض البيانات، وقد تَمَّ حذف أعمدة الحدود العمليَّة للفئات، والتعداد، والتكرار النسبي، والتكرار المئوي، والتكرار المجمَّع الهابط للأسباب التي ذُكِرت سابقاً. أي إنَّنا سنتعامل في دراستنا المقبلة مع الحدود الفعليّة للفئات فقط.

٢- لقد لاحظنا أنَّه يمكننا حساب المدى لبيانات خام بوساطة العلاقة [1,3] ، وأمَّا إذا كانت البيانات مجمَّعة في جدول توزيع تكراري بـ $k$ فئة مع $x_1$ مركز الفئة الأولى، $x_2$ مركز الفئة الثانية، ... و $x_k$ مركز الفئة الأخيرة $k$، فإنَّ المدى $\mathcal{R}$ لبيانات هذا الجدول يُحسب بوساطة العلاقة الآتية:

$$\mathcal{R} = x_k - x_1 \qquad [1,7]$$

## (١,٥) العروض البيانية لبيانات جداول التوزيع التكرارية

**Graphic representations of Frequency distribution tables**

في كثير من الحالات يكون التعامل مع جداول التوزيع التكرارية أمراً شاقاً، وخاصة إذا كان عدد الفئات كبيراً، ولذلك يُلجأ إلى عرض النتائج من خلال أشكال بيانية مناسبة يَسْهل معها فهم طبيعة وسلوك البيانات، ومن هذه الأشكال البيانية ما يلي:

### (١,٥,١) المدرَّجات التكرارية Frequency Histograms

إنَّ المدرَّج التكراري هو تمثيل بياني لبيانات مجمَّعة في فترات، ويتكوَّن هذا المدرَّج من أعمدة مستطيلة متلاصقة تُرسم فوق الفترات، وبحيث تكون مساحة كلّ مستطيل متناسبة مع التكرار وطول الفترة التي رُسم عليها.

#### (١,١,٥,١) ملاحظة

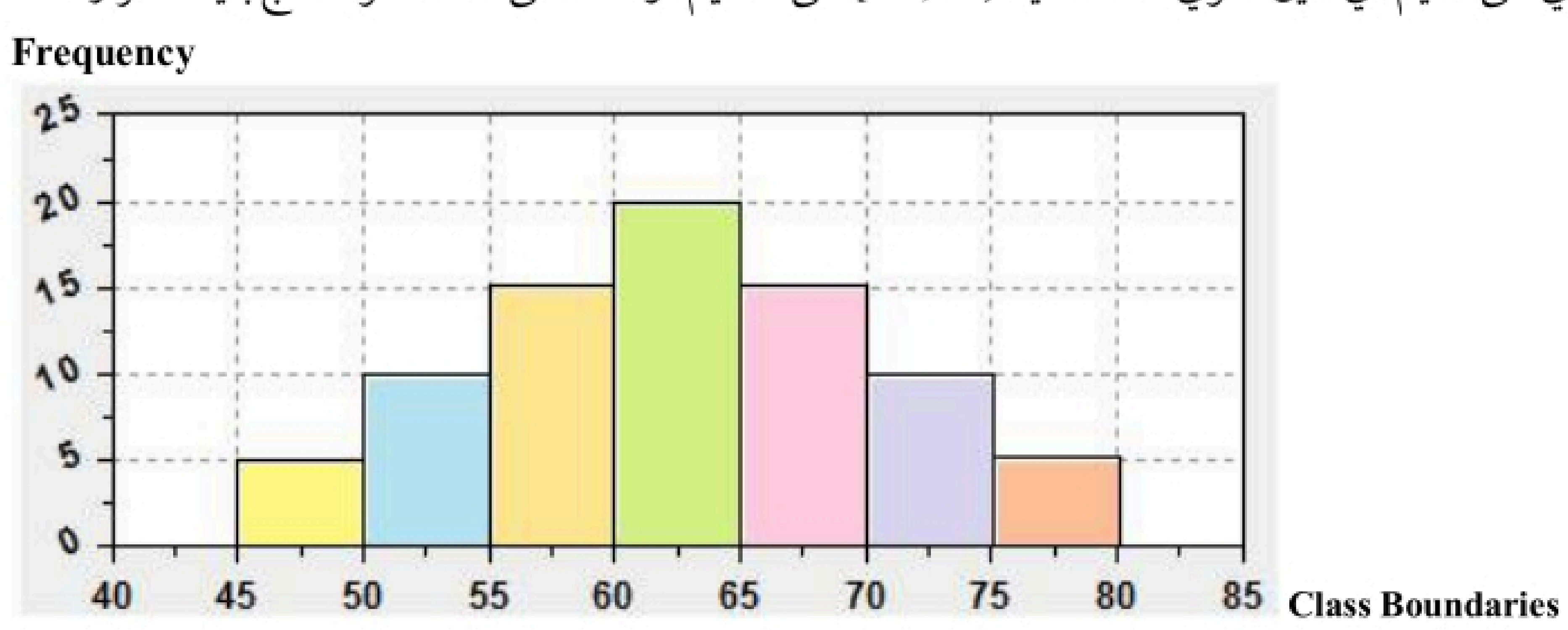
١- بما أنَّنا نتعامل مع فئات فعليّة متساوية السعة فإنَّ ارتفاع الأعمدة المستطيلة التي تُرسم فوق الفئات الفعليّة ستكون متناسبةً مع تكرار القيم الواقعة في الفئة الفعليّة التي رُسم عليها فقط.

٢- من أجل بيانات مجمَّعة في جدول توزيع تكراري مُعطى يتمّ عرض المدرَّج التكراري الموافق لذلك الجدول من خلال رسم أعمدة مستطيلة فوق الفئات الفعليّة (الفئات الفعليّة في جداول التوزيع التكرارية توافق الفترات المذكورة آنفًا) بحيث يكون ارتفاع كل مستطيل متناسباً مع تكرار الفئة التي رُسم عليها.

٣- نشير هنا إلى أنَّه إذا كانت البيانات مقدّمةً من خلال جدول توزيع تكراري بفئات منفصلة ولم يذكر فيه الحدود الفعليّة للفئات فعندئذ يجب تعيين الفئـات الفعليّة أولاً ومن ثمَّ رسم الأعمدة عليها.

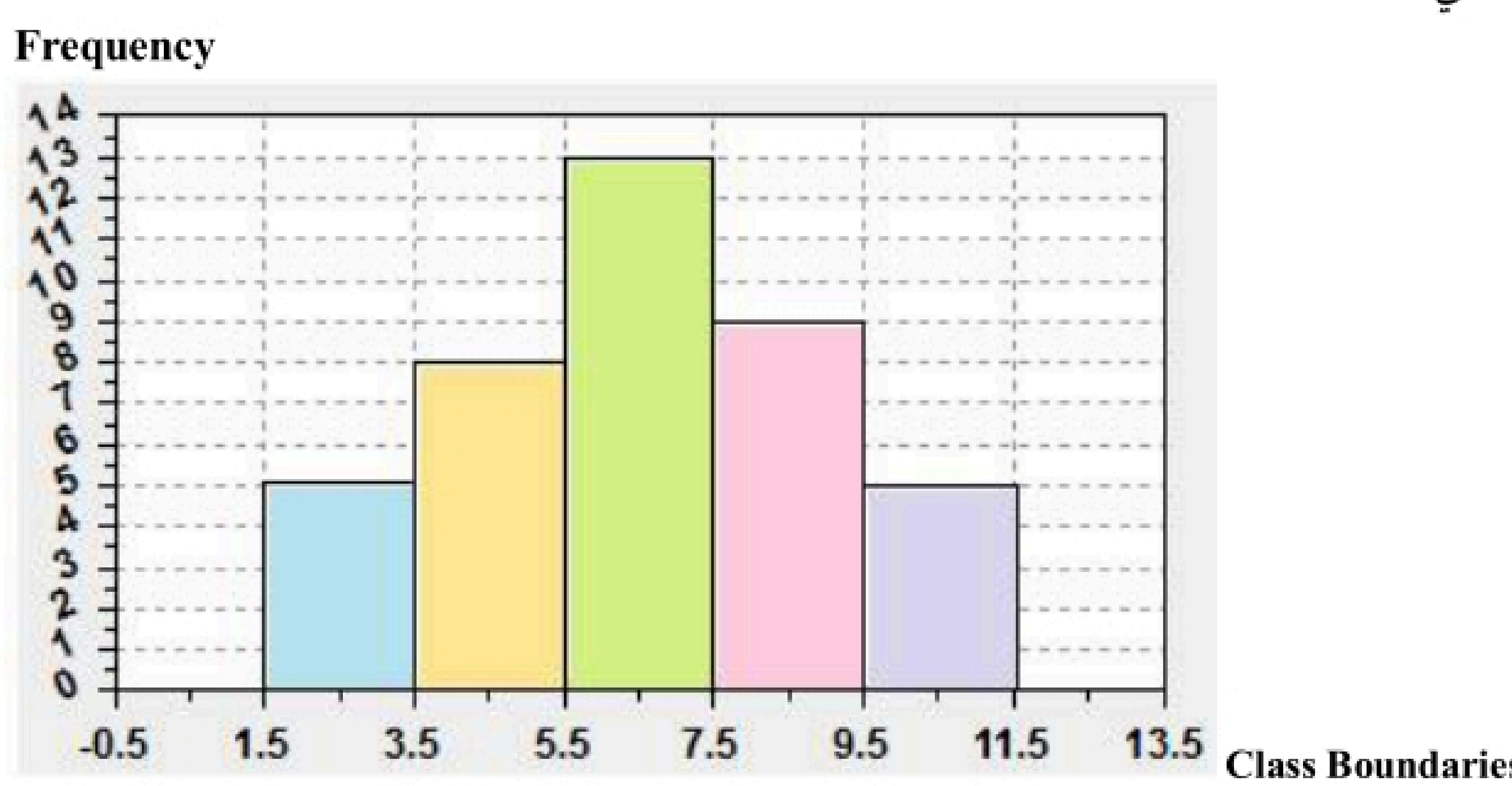
#### (١,٥,١,٢) أمثلة

١- يعرض الشكل الآتي مدرَّجاً تكرارياً لـ 80 قيمة (بيانات) مجمَّعة في سبع فئات حيث يُلاحظ أنَّ الفئة الأولى (50 , 45] تحتوي على 5 قيم، في حين تحتوي الفئة التالية (55 , 50] على 10 قيم، وهكذا على هذا النحو تستنج بقية التكرارات للفئات المتبقية.



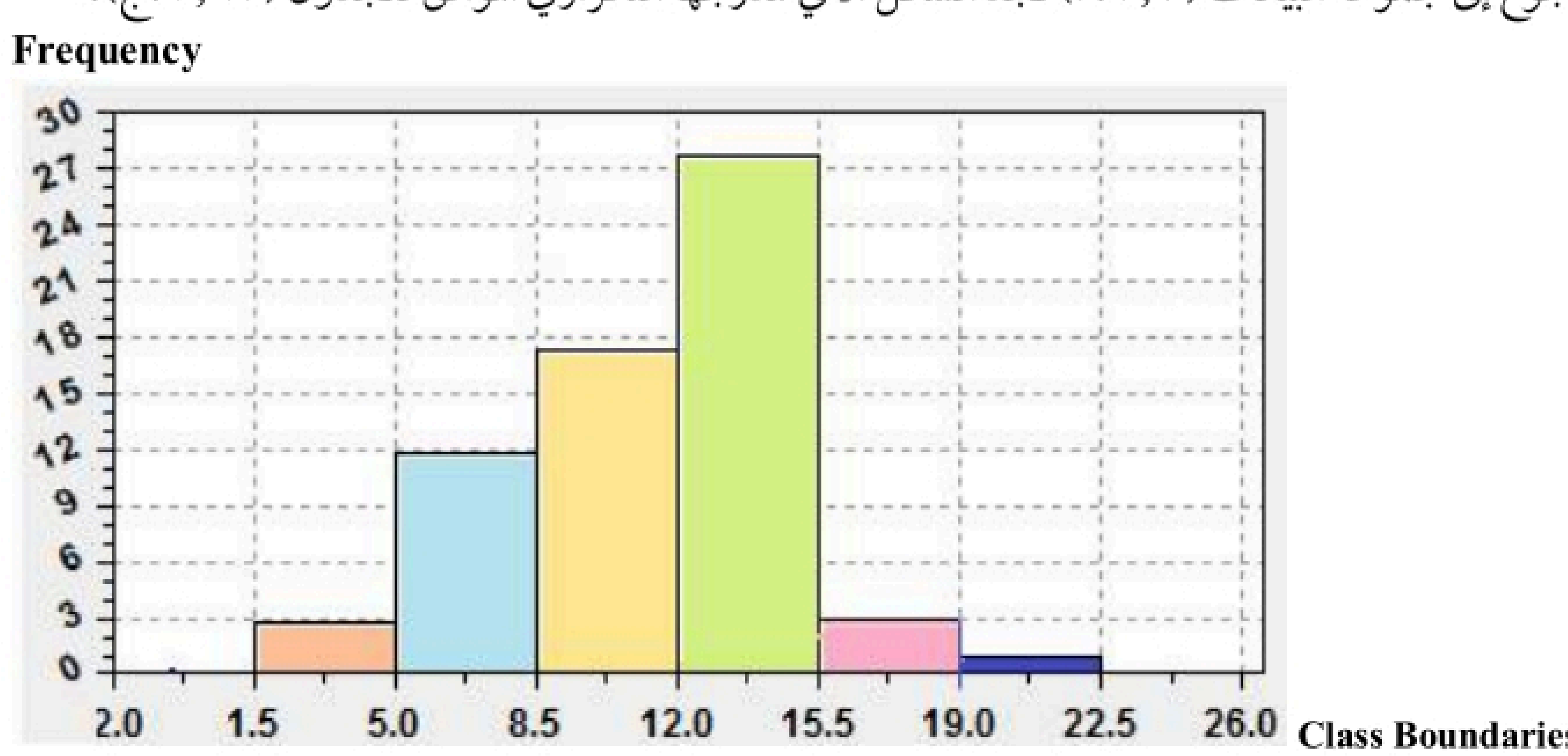

الشكل (١,١٧)

٢- بالرجوع إلى المثال / ١/ من (٢ , ٧ , ٤ , ١) نجد أنَّ المدرَّج التكراري لبيانات جدول التوزيع التكراري (١٣ , ١ .ب) له الشكل (١٨ , ١) الآتي.




الشكل (١٨ , ١)

٣- بالرجوع إلى مجموعة البيانات (٣ , ١ . أ) نجد الشكل الآتي لمدرَّجها التكراري الموافق للجدول (١٣ , ١ .ج):




الشكل (١٩ , ١)

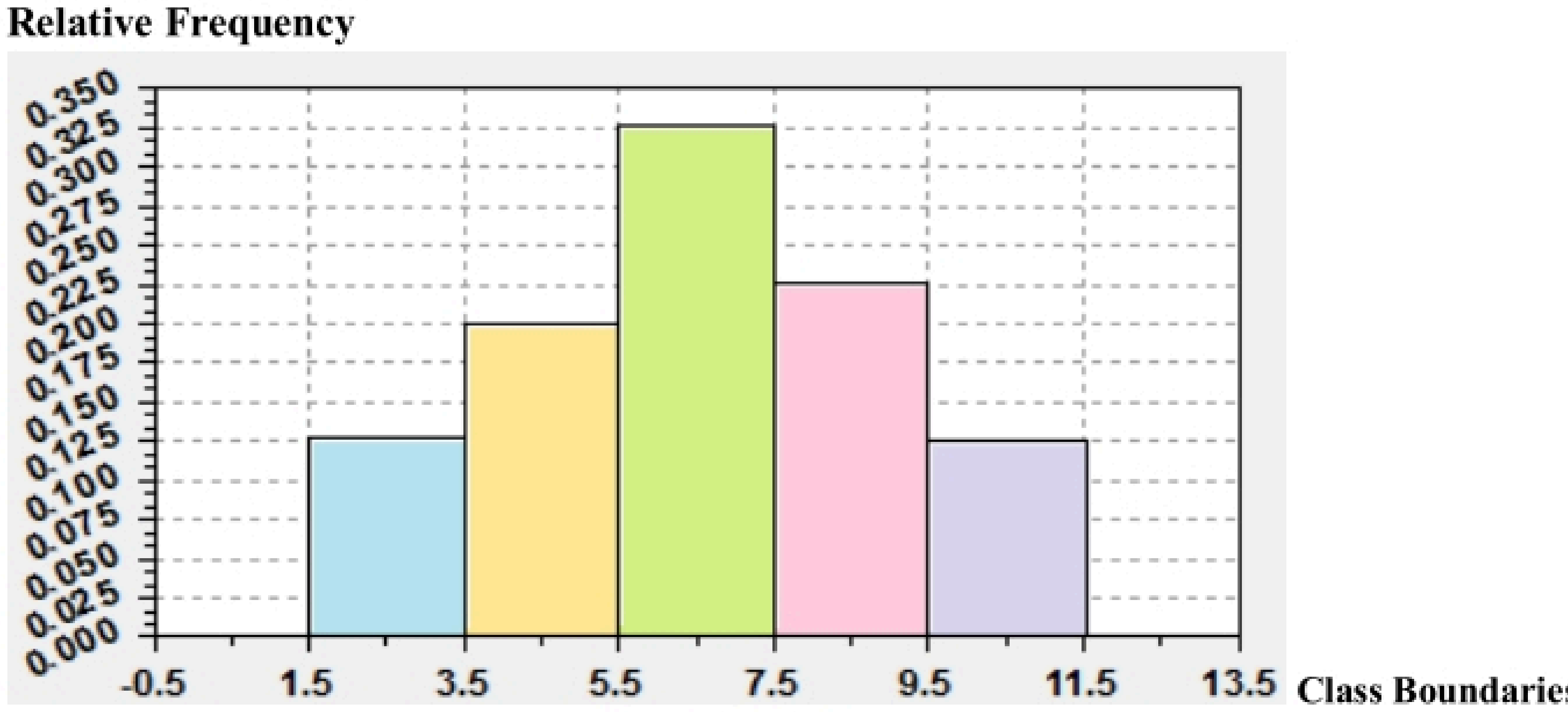
بالإضافة لما سبق يندرج تحت مسمَّى مدرَّجات التوزيع التكرارية النماذج الآتية:

### (٣ , ١ , ٥ , ١) المدرَّج التّكراري النسبيّ (أو مدرَّج نسبة التكرارات) Relative Frequency Histogram

إنَّ المدرَّج التكراري النسبي هو تمثيل بياني آخر لبيانات مجمَّعة في فترات (أو فئات فعليّة) متتالية ويتكوَّن من أعمدة مستطيلة متلاصقة أيضاً، وبحيث تكون مساحة كل مستطيل متناسبة مع التكرار النسبي للقيم الواقعة في الفترة (أو الفئة الفعليّة) التي رُسم عليها، ومن أجل مجموعة بيانات مجمَّعة مُحدَّدة سيكون شكل المدرَّج التكراري النسبي مشابهًا تماماً لشكل المدرَّج التكراري والاختلاف بينهما يكمن في وحدات القياس على محور التكرارات فقط. أمَّا من أجل بيانات مجمَّعة في جدول توزيع تكراري مُعطى فيتمّ عرض المدرَّج التكراري النسبيّ الموافق لذلك الجدول من خلال رسم أعمدة مستطيلة فوق الفئات الفعليّة بحيث يكون ارتفاع كلّ مستطيل متناسباً مع التكرار النسبي للفئة التي رُسم عليها.

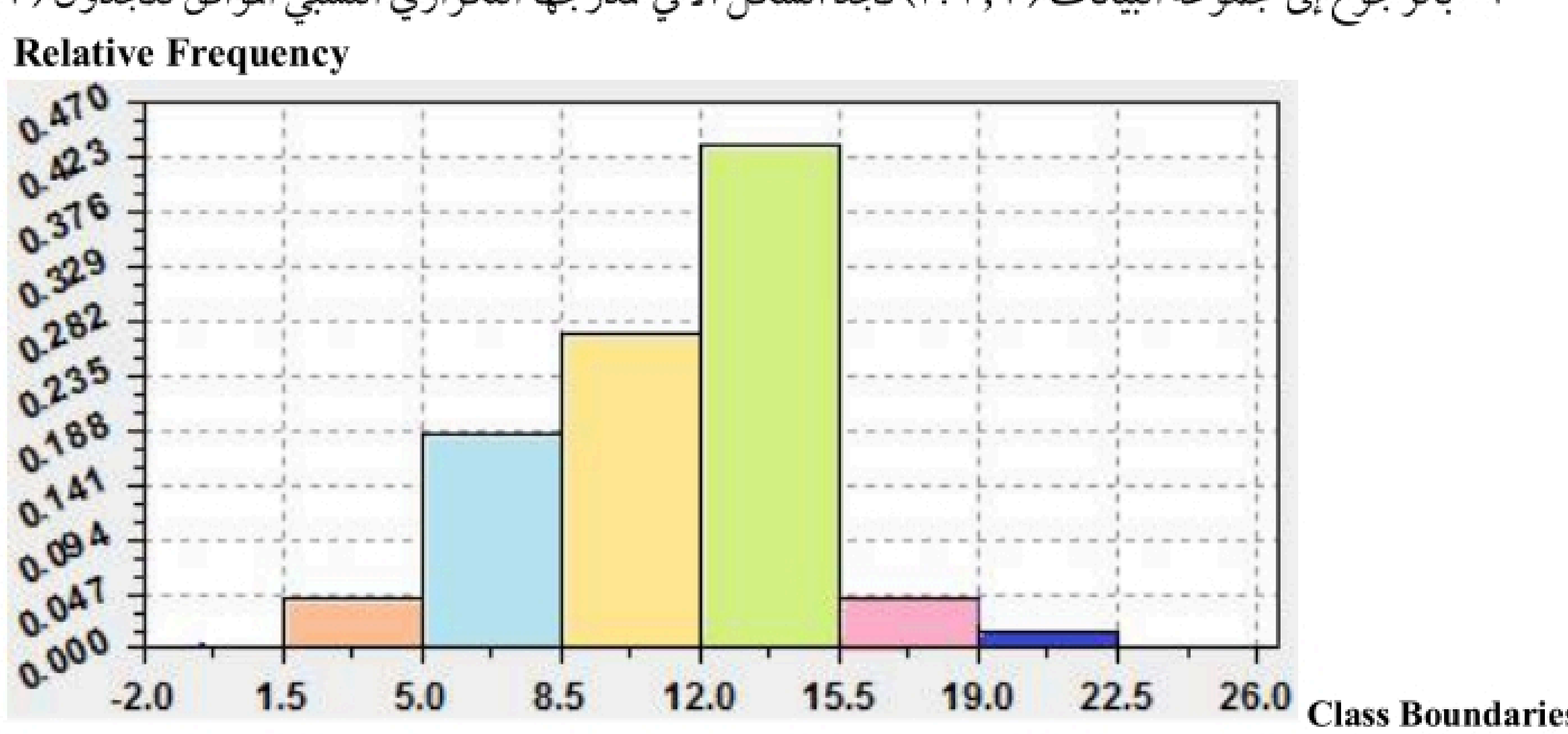
(١,٥,١,٣,١) أمثلة

١- بالرجوع إلى بيانات المثال / ١/ من (٢,٧,٤,١) نجد لمدَّرجها التكراري النسبيّ الشكل الآتي:




الشكل (١,٢٠)

٢- بالرجوع إلى مجموعة البيانات (٣,١.أ) نجد الشكل الآتي لمدَّرجها التكراري النسبيّ الموافق للجدول (١٣,١.ج):



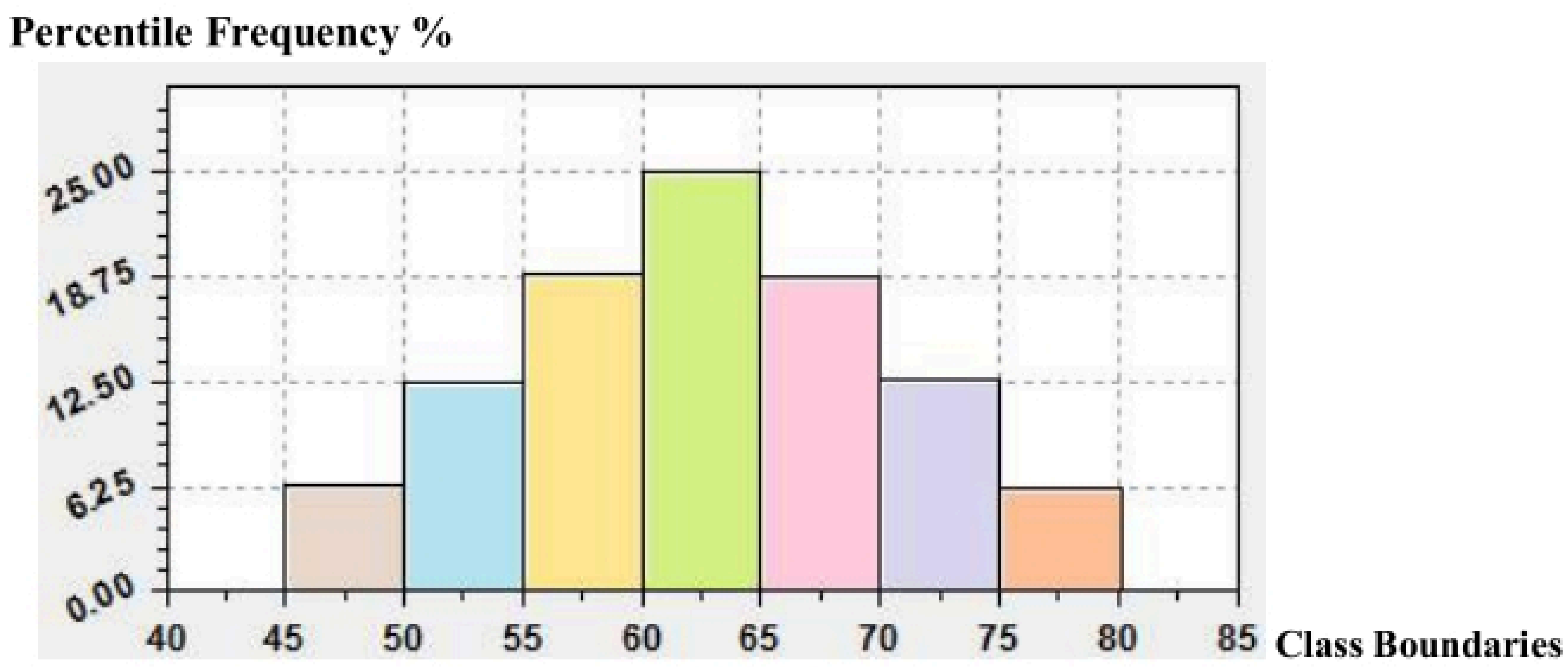

الشكل (١,٢١)

### (١,٥,١,٤) المدَّرج التكراري المئوي (أو مدرَّج تكرارات النسبة المئوية) Percentile Frequency Histogram

إنَّ المدرَّج التكراري المئوي هو تمثيل بياني آخر لبيانات مجمَّعة في فئات متتالية أيضاً، ويتكوَّن من أعمدة مستطيلة متلاصقة، وبحيث تكون مساحة كل مستطيل متناسبة مع التكرار المئوي للقيم الواقعة في الفئة التي رُسم عليها، ولذلك من أجل مجموعة بيانات مجمَّعة مُحدَّدة يكون شكل المدرَّج التكراري المئوي مشابهًا تماماً لشكل المدرَّج التكراري والتكرار النسبي، والاختلاف بينهما يكمن في وحدات القياس على محور التكرارات فقط.

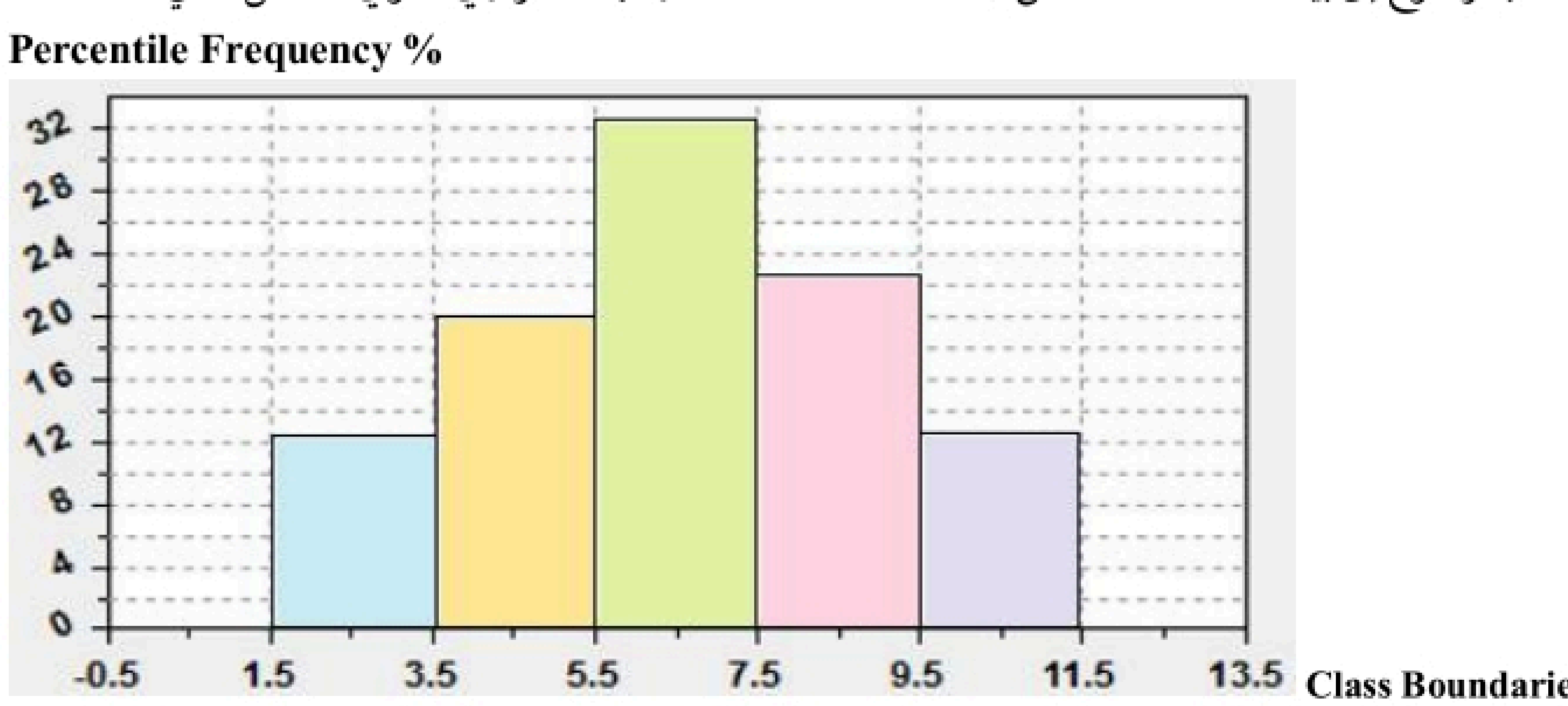
(١,٥,١,٥) أمثلة

١- الشكل الآتي يعرض لنا المدرَّج التكراري المئوي للبيانات الـــ 80 الممثَّلة بالمدرَّج التكراري المُعطى بالشكل (١,١٥)، حيث يُلاحظ أنَّ النسبة المئوية للبيانات في الفترة الأولى (50 , 45] يساوي % 100×(5/80) في حين أنّنا نجد أنَّ النسبة المئوية للبيانات في الفترة التالية (55 , 50] هي % 100×(10/80)، وعلى هذا النحو تُستنتج بقية التكرارات النسبية في بقية الفترات.



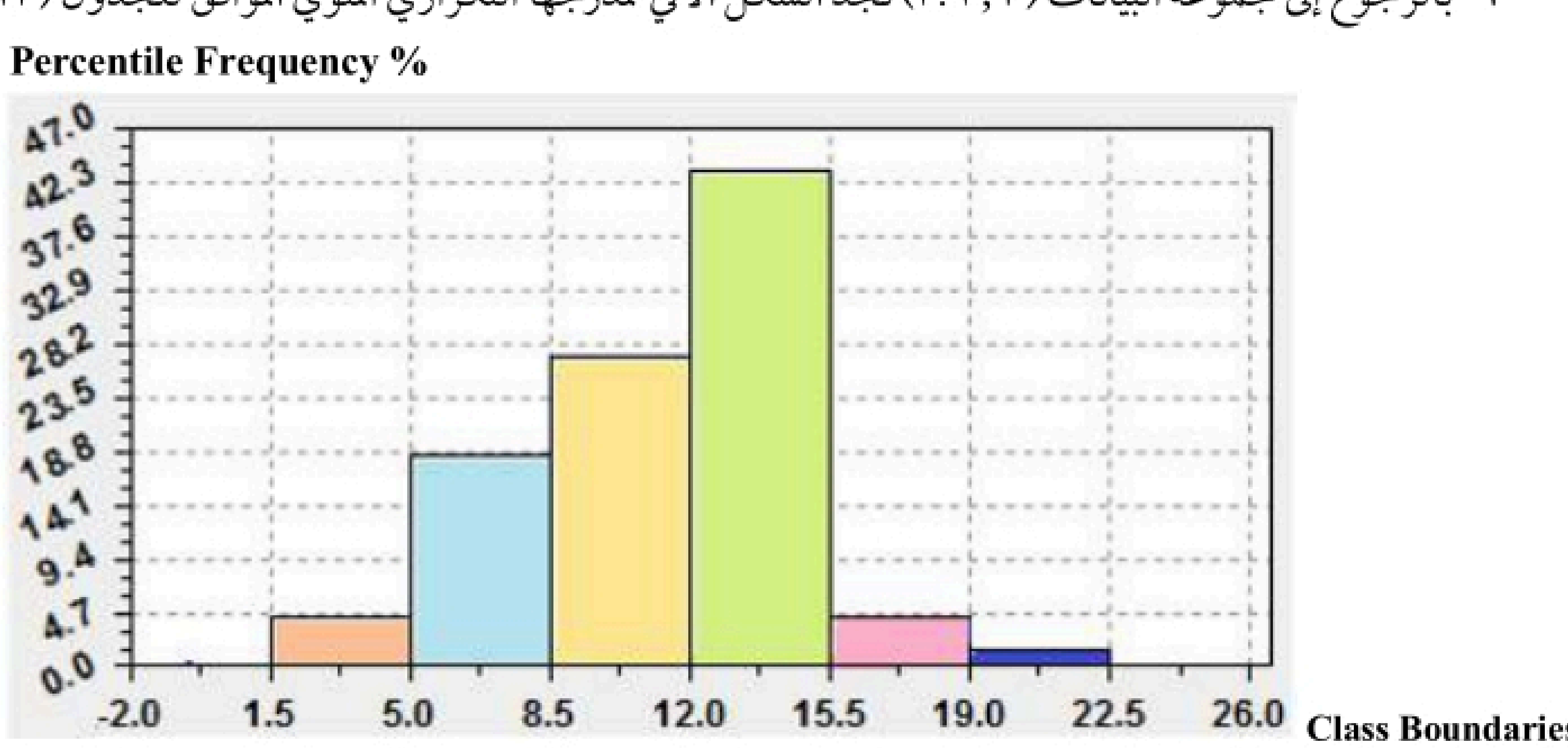

الشكل (١,٢٢)

٢- بالرجوع إلى بيانات المثال / ١/ من (١,٤,٧,٢) نجد لمدرَّجها التكراري المئويّ الشكل الآتي:




الشكل (١,٢٣)

٣- بالرجوع إلى مجموعة البيانات (١,٣ . أ) نجد الشكل الآتي لمدرَّجها التكراري المئويّ الموافق للجدول (١٣ , ١ .ج):




الشكل (١,٢٤)

٤ - لتكن لدينا مجموعة البيانات الآتية التي تمثّل الطول لـ 58 شخصاً بالغاً مقدّرةً بالسنتيمتر:

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 171 | 159 | 170 | 173 | 155 | 169 | 167 | 182 | 171 | 203 | 155 | 194 | 167 | 186 | 208 |
| 190 | 167 | 167 | 152 | 184 | 168 | 186 | 172 | 170 | 153 | 149 | 185 | 199 | 181 | 171 |
| 159 | 170 | 172 | 181 | 191 | 150 | 166 | 148 | 177 | 180 | 179 | 180 | 172 | 165 | 193 |
| 145 | 175 | 172 | 165 | 184 | 185 | 196 | 154 | 140 | 173 | 173 | 187 | 160 | | |

ولنقم بصب هذه البيانات في جدول توزيع تكراري بحيث تكون سعة الفئات الفعليّة فيه تساوي $\mathbf{C} = 10$، فنجد أنَّ عدد الفئات المطلوبة لهذا الجدول يساوي:

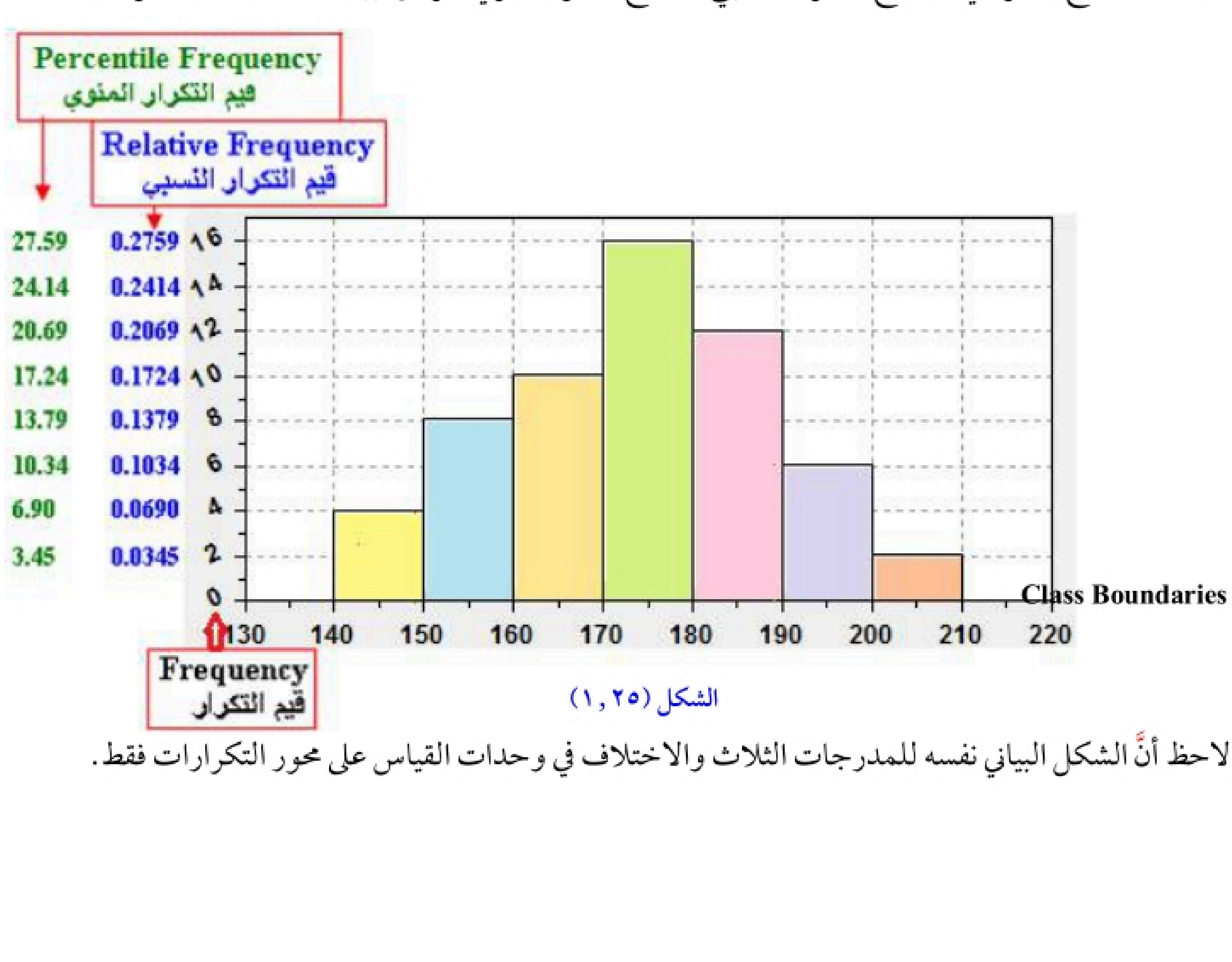
$$k = \left\lceil \frac{\mathcal{R}}{\mathbf{C}} \right\rceil = \left\lceil \frac{(208-140)+1}{10} \right\rceil = \lceil 6.9 \rceil = 7$$

ومن ثَمَّ يكون لجدول التوزيع التكراري للبيانات المُعطاة العرض الآتي:

الجدول (١,١٥)

| رقم الفئة | الحدود الفعليّة للفئة | تكرار الفئة | التكرار النسبي للفئة | التكرار المئوي للفئة | التكرار المتجمّع الصاعد الفئة |
|---|---|---|---|---|---|
| 1 | 140 → 150 | 4 | 0.0690 | 6.90 % | 4 |
| 2 | 150 → 160 | 8 | 0.1379 | 13.79 % | 12 |
| 3 | 160 → 170 | 10 | 0.1724 | 17.24 % | 22 |
| 4 | 170 → 180 | 16 | 0.2759 | 27.59 % | 38 |
| 5 | 180 → 190 | 12 | 0.2069 | 20.69 % | 50 |
| 6 | 190 → 200 | 6 | 0.1034 | 10.34 % | 56 |
| 7 | 200 → 210 | 2 | 0.0345 | 3.45 % | 58 |
| Total | ------------- | 58 | 1 | 100 % | المجموع |

فنجد أنَّ المدرَّج التكراري، المدرج التكرار النسبي ومدرج التكرار المئوي الموافق لبيانات هذا الجدول يعرضها لنا الشكل الآتي:




الشكل (١,٢٥)

لاحظ أنَّ الشكل البياني نفسه للمدرجات الثلاث والاختلاف في وحدات القياس على محور التكرارات فقط.

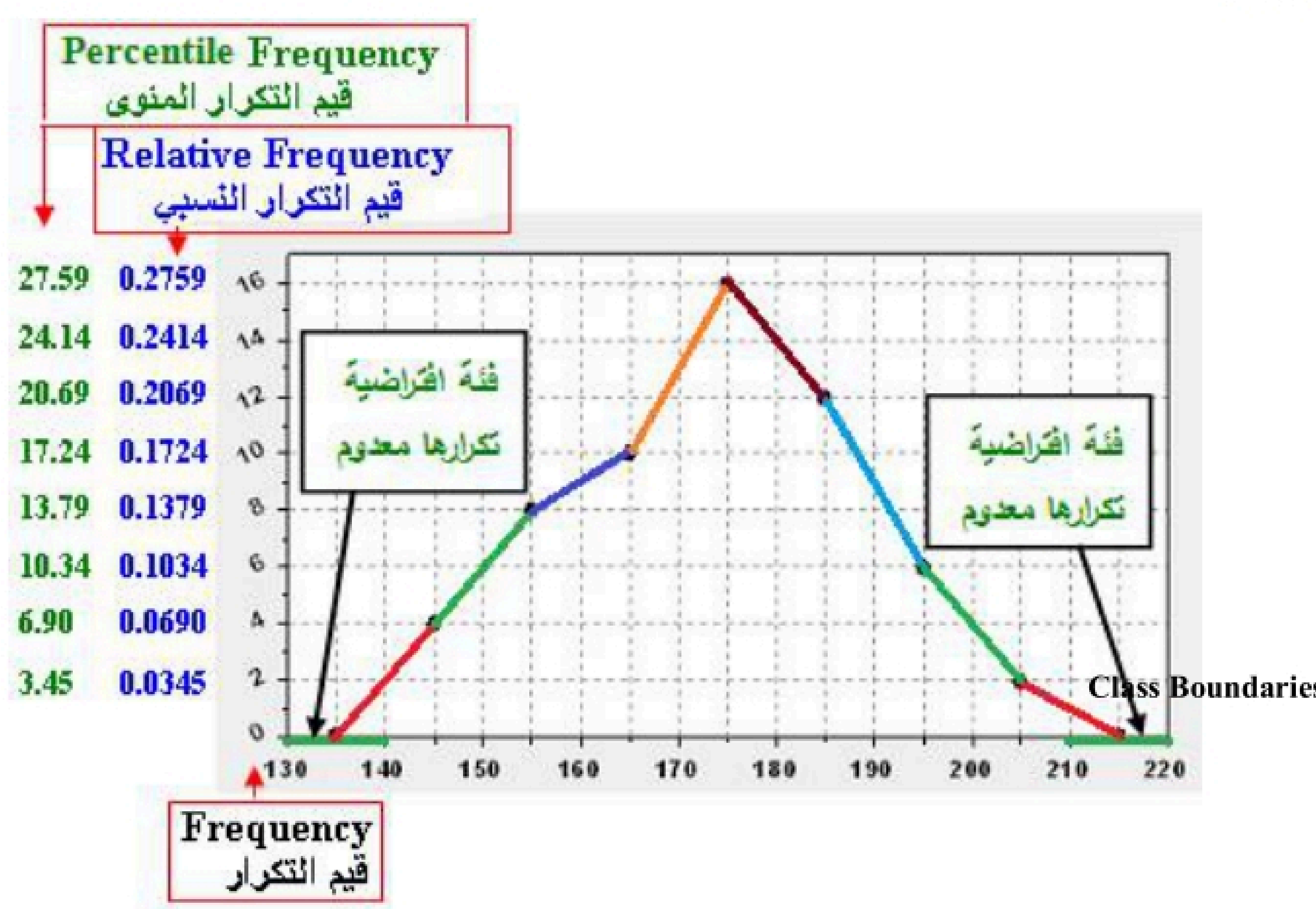
### (٢ , ٥ , ١) المضلَّعات التكرارية Frequency Polygons:

يُنظر في بعض الحالات إلى أنَّ كل فئة يمكن تمثيلها بمركزها، وهـذا يقودنا إلى إنشاء ما يسمَّى بـ المضلَّع التكراري (المضلَّع التكراري النسبي، المضلَّع التكراري المئوي) الذي يُشكَّل من خلال الوصل بقطع مستقيمة بين تلك النقاط التي إحداثياتها على محور الفئات هي مراكز الفئات، وأمّا إحداثياتها على محور التكرارات (التكرارات النسبية، التكرارات المئوية) فهي قيم التكرارات (التكرارات النسبية، التكرارات المئوية) المقابلة لتلك الفئات، وبعد ذلك إغلاق هذا المضلَّع إلى محور الفئات من خلال وصل بداية المضلَّع النّاتج إلى مركز فئة وهمية سابقة لأول فئة تكرارها (تكرارها النسبي، تكرارها المئوي) معدوم، وبعد ذلك وصل نهاية المضلَّع النّاتج إلى مركز فئة وهمية لاحقة بآخر فئة تكرارها (تكرارها النسبي، تكرارها المئوي) معدوم.

إنَّ الشكل البياني للمضلَّعات التكرارية الثلاث ( المضلَّع التكراري، والمضلَّع التكراري النسبي، والمضلَّع التكراري المئوي) متشابهة تماماً، والاختلاف فيما بينها يكمن في وحدات القياس على محور التكرارات فقط.

### (١ , ٢ , ٥ , ١) مثال:

بالرجوع إلى بيانات جدول التوزيع التكراري (١٥ , ١) السّابق فإنَّنا نجد المضلَّع التكراري ( المضلَّع التكراري النسبيّ، المضلَّع التكراري المئويّ) للبيانات كما في الشّكل الآتي (حيث نلاحظ أنَّ الشكل البياني نفسه للمضلَّعات الثلاث والاختلاف في وحدات القياس على محور التكرارات فقط).




الشكل (٢٦ , ١). المضلَّع التكراري لبيانات الجدول (١٥ , ١).

### (٢ , ٢ , ٥ , ١) ملاحظة

نشير هنا إلى أنَّ بعض المراجع تقوم بإغلاق الطرف الأيسر للمضلَّع إلى بداية الفئة الأولى، وإغلاق الطرف الأيمن للمضلَّع إلى نهاية الفئة الأخيرة، ويمكن استخدام مثل هذا الأسلوب في الإغلاق من أجل بعض الحالات الخاصّة (انظر الشكل (٢٦ , ١.ب) في التمرين القادم / ١/ من (٥ , ٥ , ١)).

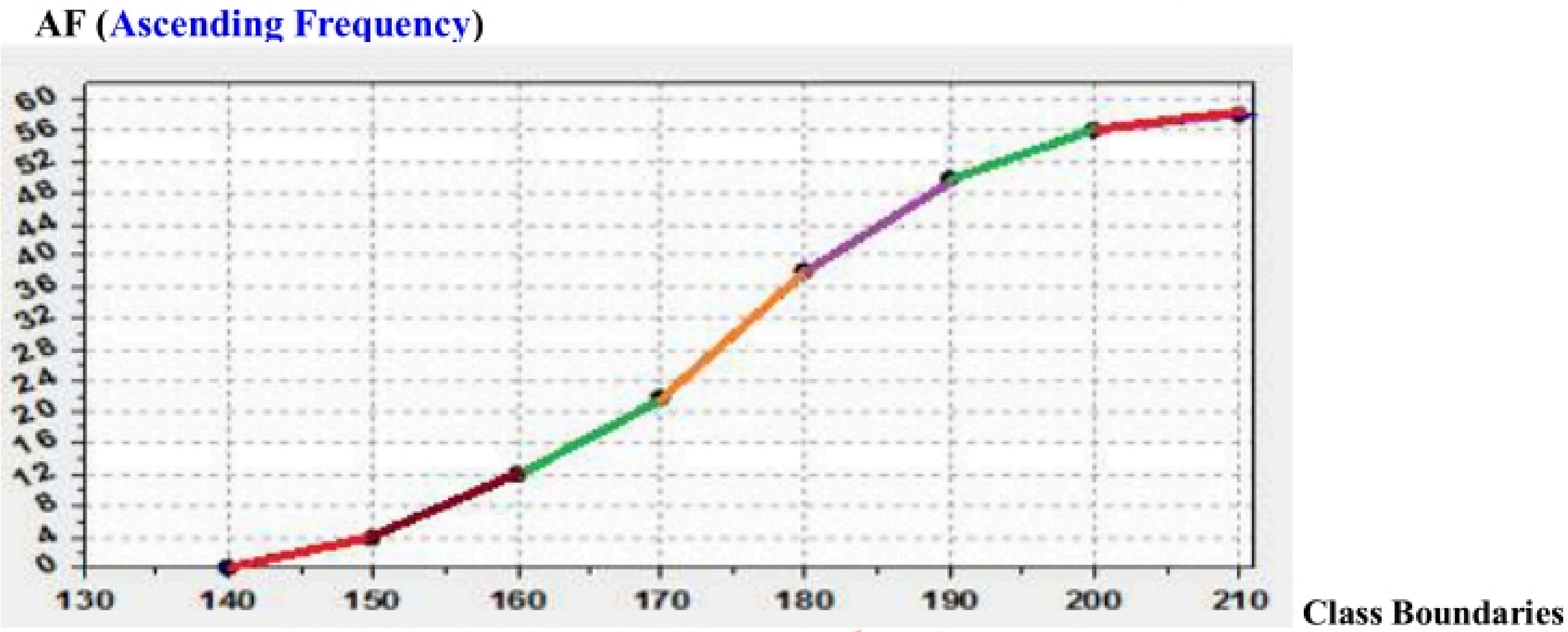
### (٣, ٥, ١) المضلَّعات التكرارية التراكمية Cumulative Frequency Paragons

من أجل هذا النوع من المضلَّعات التكرارية نميِّز بين نوعين من المضلَّعات وهما:

#### (١, ٣, ٥, ١) مضلَّع التكرار المتجمِّع الصاعد Ascending Cumulative Frequency Paragon

إنَّ مضلَّع التكرار المتجمِّع الصاعد (ويدعى أقل من القوس القوطي أيضاً Less than Ogive) ينشأ هذا المضلَّع من خلال الوصل بقطع مستقيمة بين النقاط الناشئة عن الحدود العليا للفئات الفعليّة والقيم الموافقة لها من التكرار المتجمِّع الصاعد، ومن ثمَّ إغلاق بدايته مع بداية الفئة الأولى.

على سبيل المثال لو أخذنا بيانات جدول التوزيع التكراري (١٥, ١) السابقة لوجدنا أنَّ مضلَّع التكرار المتجمّع الصاعد لتلك البيانات له الشكل الآتي:

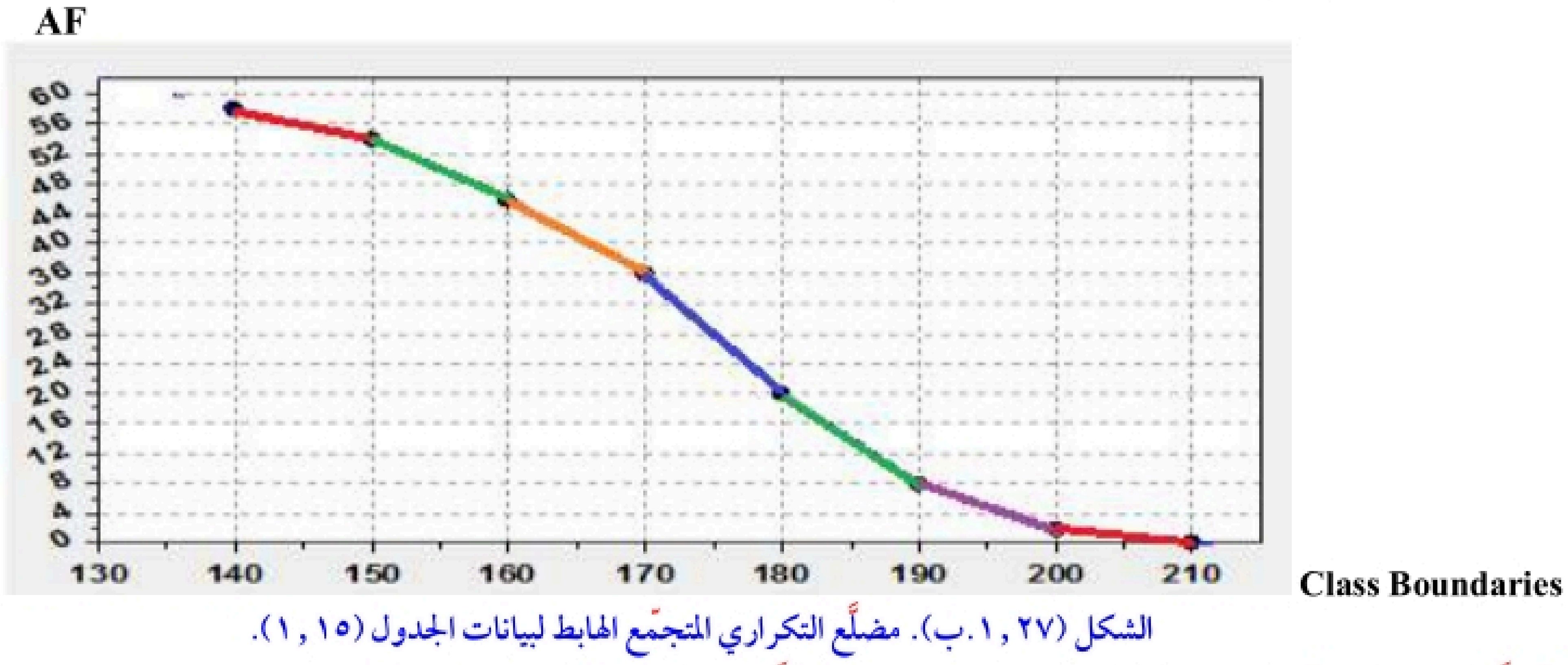



الشكل (٢٧, ١ .أ). مضلَّع التكرار المتجمّع الصاعد لبيانات الجدول (١٥, ١).

#### (٢, ٣, ٥, ١) مضلَّع التكرار المتجمِّع الهابط Descending Cumulative Frequency Polygon

إنَّ مضلَّع التكرار المتجمِّع الهابط (ويدعى أكثر من القوس القوطي أيضاً More than Ogive) ينشأ هذا المضلَّع من الوصل بقطع مستقيمة بين النقاط الناشئة عن الحدود الدنيا للفئات الفعليّة والقيم الموافقة لها من التكرار المتجمِّع الهابط، ومن ثمَّ إغلاقه مع نهاية الفئة الأخيرة، وهذا النوع من المضلَّعات قليل الاستخدام.

الشكل الآتي يعرض لنا مضلَّع التكرار المتجمّع الهابط لبيانات الجدول (١٥, ١).




الشكل (٢٧, ١ .ب). مضلَّع التكراري المتجمّع الهابط لبيانات الجدول (١٥, ١).

علماً أنَّ التكرار المتجمّع الهابط لبيانات الجدول (١٥, ١) موضَّح في الجدول الآتي:

الجدول (١٦, ١).

| رقم الفئة | الحدودلفعليّة للفئة | تكرار الفئة | التكرار المتجمّع الهابط للفئة |
|---|---|---|---|
| 1 | 140 → 150 | 4 | 58 |
| 2 | 150 → 160 | 8 | 54 |
| 3 | 160 → 170 | 10 | 46 |
| 4 | 170 → 180 | 16 | 36 |
| 5 | 180 → 190 | 12 | 20 |
| 6 | 190 → 200 | 6 | 8 |
| 7 | 200 → 210 | 2 | 2 |
| Total | ------------- | 58 | ------------- |

إنَّ كلاً من عبارتي " **أقل من القوس القوطي** " و " **أكثر من القوس القوطي** " مرتبطةً بطريقة أخذ التكرارات التراكمية عند بناء المضلّع التكراري التراكمي، فعند أخذ التكرارات التراكمية التي أقل من الحدّ الأعلى للفئة الفعليّة نحصل على أقل من القوس القوطي، وفي حال أخذ التكرارات التراكمية التيأكثر من الحدّ الأدنى للفئة الفعليّة نحصل على أكثر من القوس القوطي.

أخيراً ننوِّه إلى أنَّ كلاً من القوسين الناشئين عن تقاطع مضلَّعي التكراري المتجمّع الصاعد والهابط يُدعى **قوساً قوطياً** Ogive تشبيهاً بالقوس القوطي المستخدم في بعض الأبنية القديمة والكاتدرائيات، والشكل الآتي يوضِّح ذلك.


الشكل (٢٨, ١) القوسين القوطيين الناشئين عن تقاطع مضلَّعي التكراري المتجمّع الصاعد والهابط لبيانات الجدول (١٥, ١)

### (٤, ٥, ١) المنحنيات التكرارية Frequency Curves

هي منحنيات تنتج عن المضلَّعات التكرارية (بأنواعها السابقة) وبطريقة مماثلَة تماماً لما تَّم شرحه في الفقرة (٥, ٣, ١)، ولذلك نادراً ما يستخدم هذا النوع من المنحنيات في المجالات التجريبية (العمليّة).

### (٥, ٥, ١) تمارين

١- لتكن لدينـا البيانات الآتية التي تمثّل درجات الاختبار النهائي لـ 99 طالبًا في مقرّر الإحصاء (الدرجة من 60).

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 48 | 40 | 15 | 30 | 27 | 29 | 25 | 37 | 20 | 34 | 56 | 30 | 46 | 46 | 37 | 36 | 26 |
| 38 | 9 | 0 | 38 | 18 | 29 | 37 | 31 | 27 | 30 | 20 | 32 | 35 | 40 | 21 | 25 | 36 |
| 38 | 56 | 36 | 40 | 46 | 39 | 21 | 5 | 39 | 32 | 47 | 33 | 34 | 39 | 54 | 51 | 19 |
| 26 | 47 | 47 | 33 | 32 | 20 | 36 | 20 | 45 | 48 | 36 | 60 | 39 | 36 | 18 | 46 | 38 |
| 44 | 9 | 45 | 20 | 8 | 40 | 12 | 27 | 46 | 47 | 20 | 46 | 12 | 29 | 57 | 58 | 39 |
| 4 | 18 | 57 | 49 | 50 | 52 | 39 | 34 | 33 | 34 | 55 | 55 | 35 | 60 | | | |

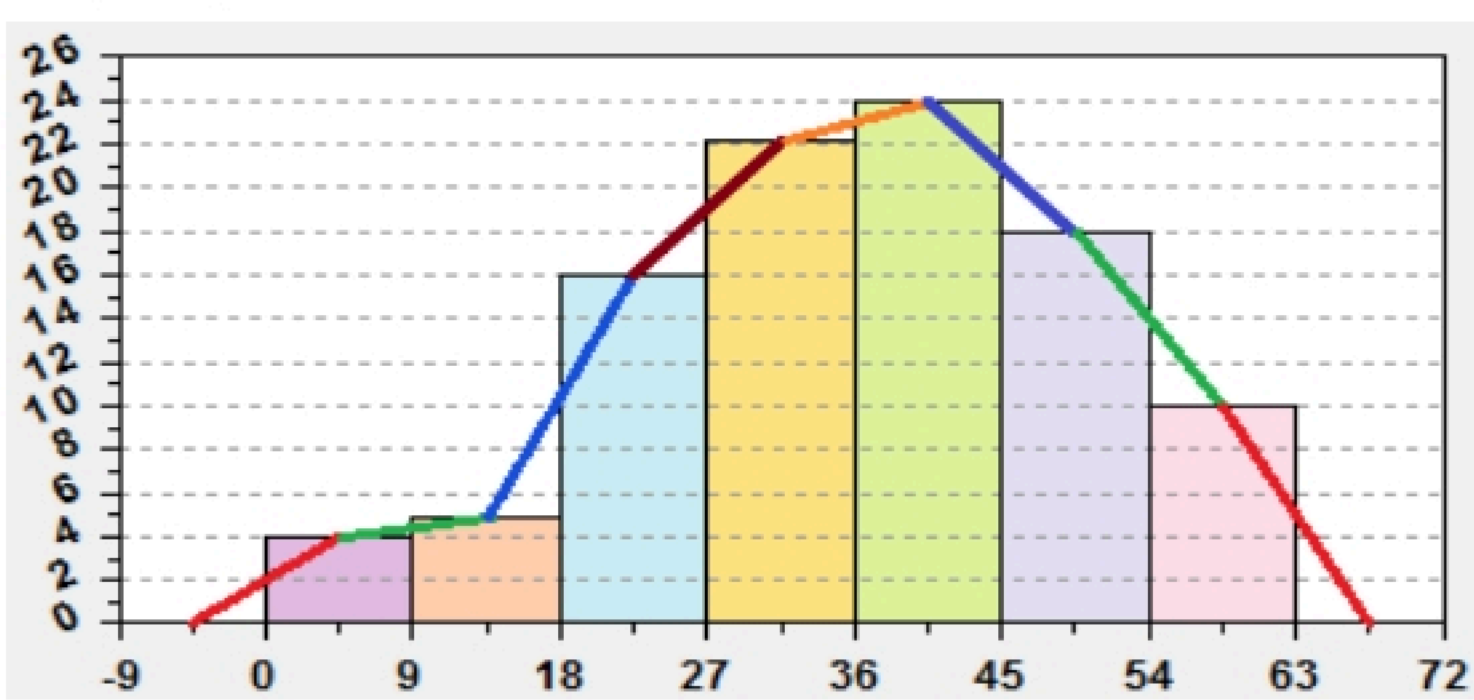
ولنقم بصبِّ هذه البيانات في جدول توزيع تكراري تكون فيه سعة الفئة الفعليّة تساوي $C = 9$، ومن ثَّم لنقم بتقديم العروض البيانية: المدرَّج التكراري، المضلّع التكراري ومضلَّع التكرار المتجمِّع.

**التنفيذ**: إنَّ جدول التوزيع التكراري للبيانات المُعطاة له العرض الآتي في الجدول (١٧ , ١):

الجدول (١٧ , ١)

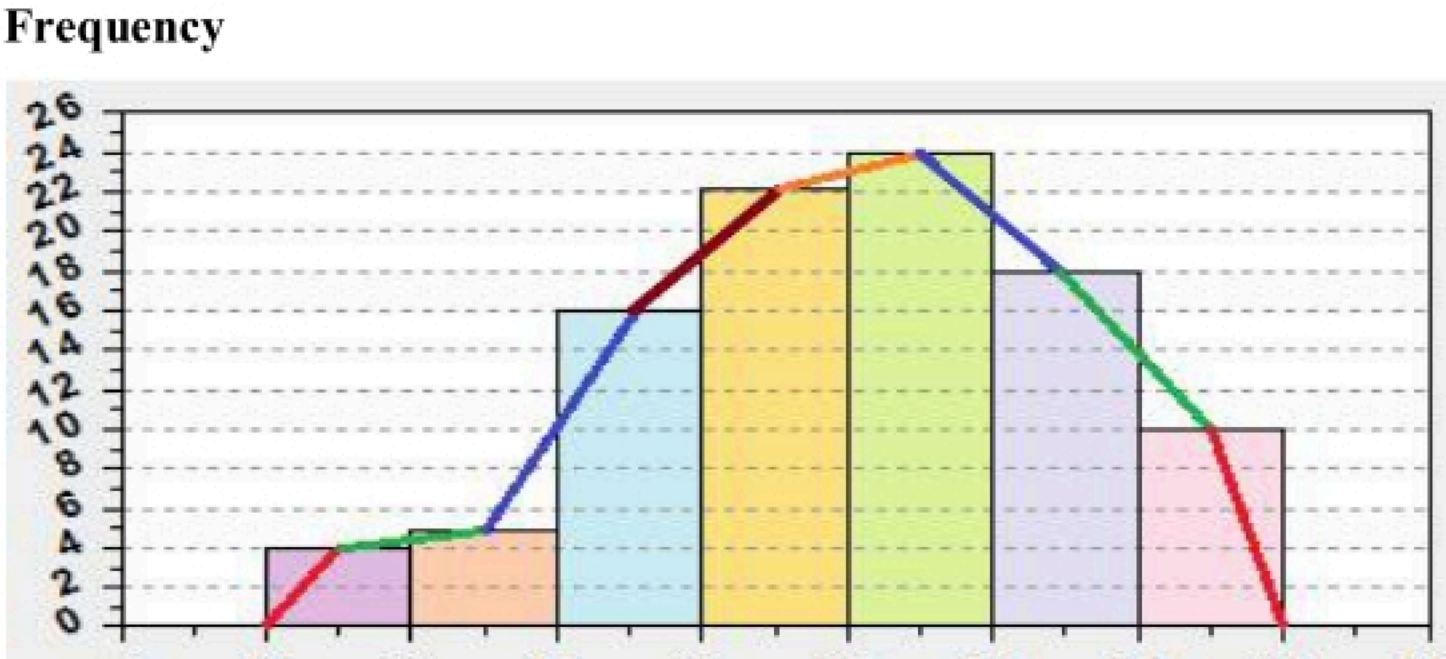
| رقم الفئة | الحدود الفعليّة للفئة | مركز الفئة | تكرار الفئة | التكرار المتجمّع الهابط للفئة |
|---|---|---|---|---|
| 1 | $0 \rightarrow 9$ | 4.5 | 4 | 4 |
| 2 | $9 \rightarrow 18$ | 13.5 | 5 | 9 |
| 3 | $18 \rightarrow 27$ | 22.5 | 16 | 25 |
| 4 | $27 \rightarrow 36$ | 31.5 | 22 | 47 |
| 5 | $36 \rightarrow 45$ | 40.5 | 24 | 71 |
| 6 | $45 \rightarrow 54$ | 49.5 | 18 | 89 |
| 7 | $54 \rightarrow 63$ | 58.5 | 10 | 99 |
| Total | ------------ | ---- | 99 | ------------- |

ومن ثمَّ يكون للمدرَّج والمضلَّع التكراري لبيانات الجدول السّابق العرض الآتي:




الشكل (٢٩ , ١.أ)

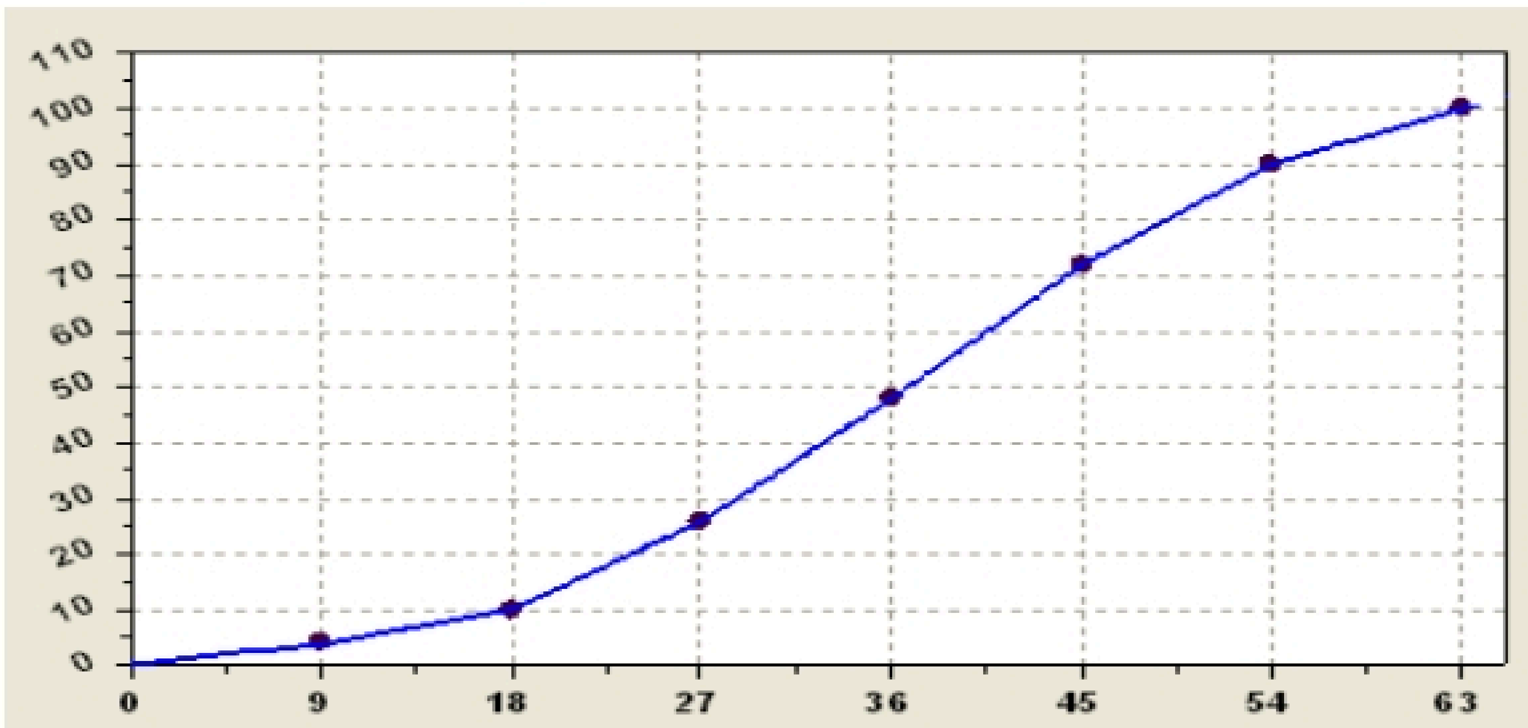
لقد ذكرنا سابقاً أنَّ بعض المراجع تغلق المضلَّع التكراري إلى طرفي الفئة الأولى والأخيرة وهذا يوافق حالة مثالنا هذا، إذ إنَّه لا توجد درجة اختبار طالب دون الصفر ولا أكبر من 60، فيجوز في مثل هذه الحالة الإغلاق إلى بداية الفئة الأولى ونهاية الفئة الأخيرة (أو إلى القيمة 60)، فيصبح للمضلَّع التكراري السابق **العرض الآتي**:




الشكل (٢٩ , ١.ب)

وأخيراً يكون لمضلَّع التكرار المتجمّع الصاعد العرض الآتي:

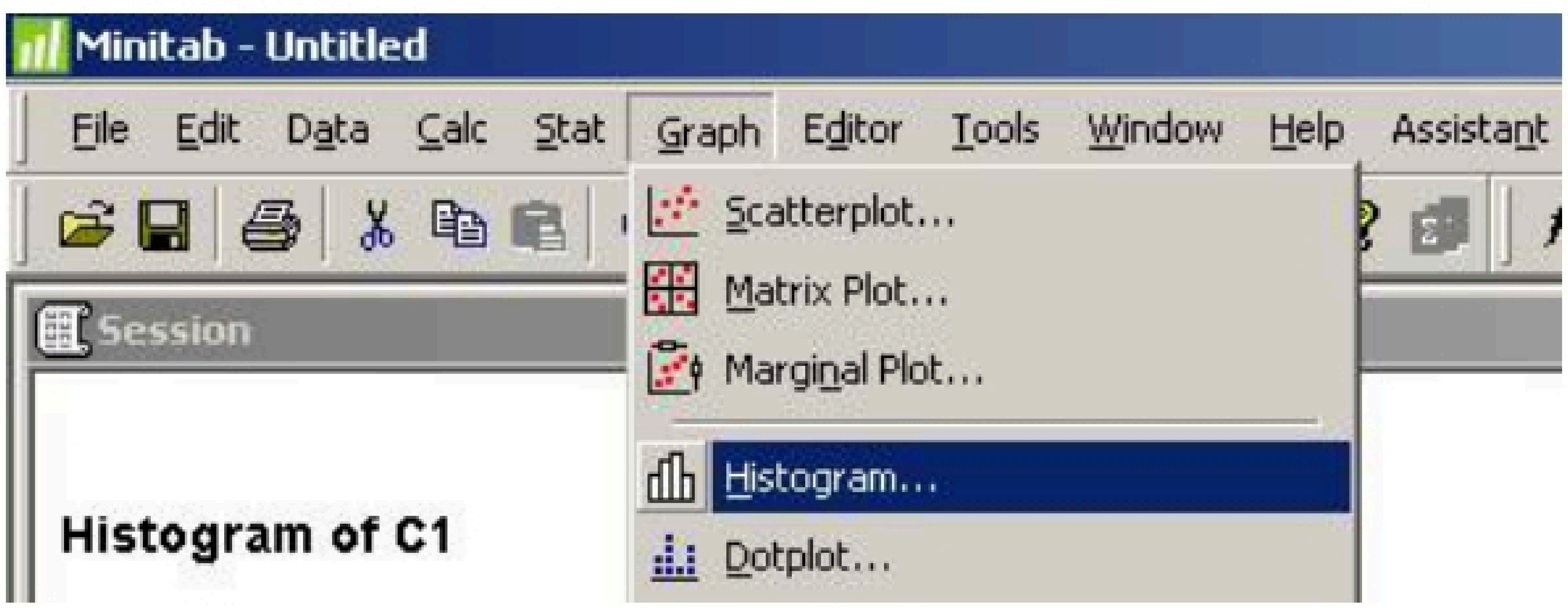
**AF (Ascending Frequency)**



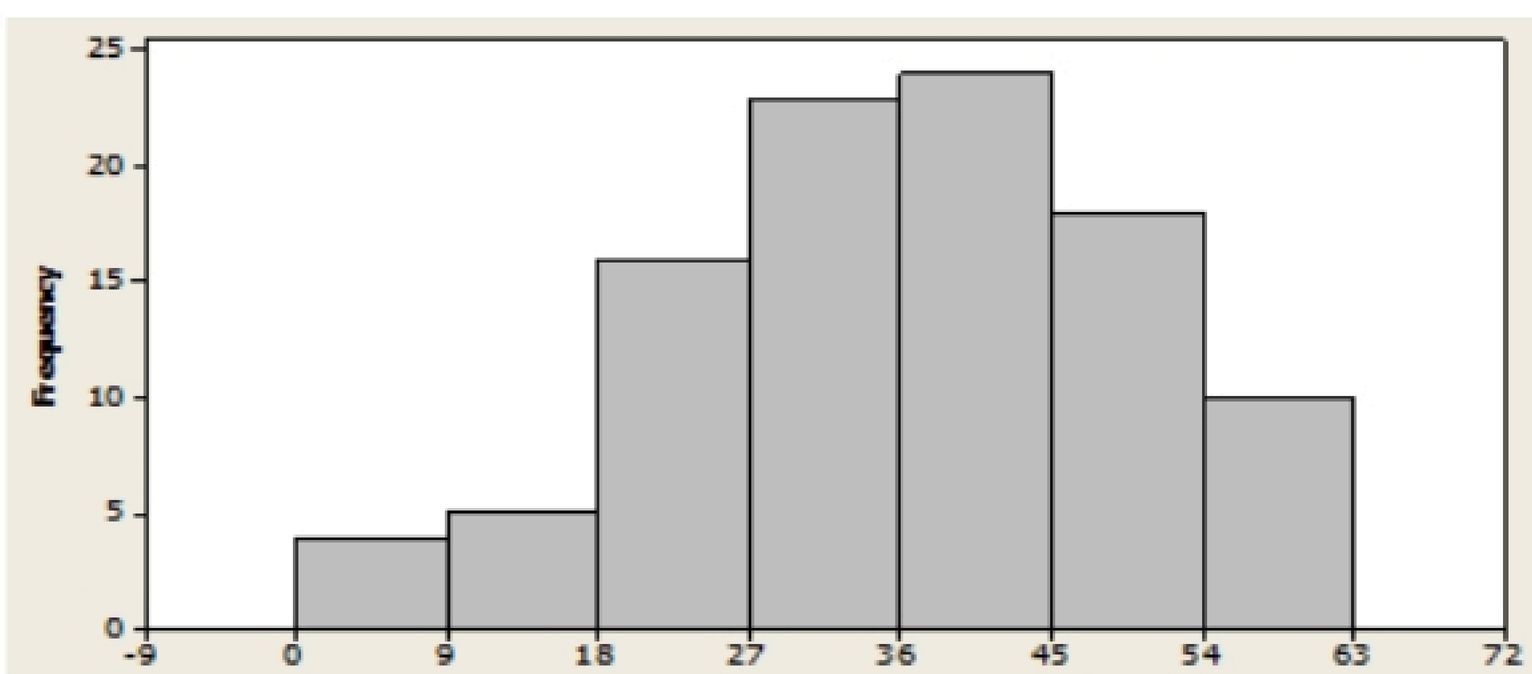

**Class Boundaries**

الشكل (٢٩, ١.ج)

ومن الممكن إنجاز المدرَّج التّكراري لبيانات جدول توزيع تكراري باستخدام برنامج Minitab من فقرة Graph في شريط الأوامر ومن ثمَّ اختيار Histogram.




فعلى سبيل المثال للحصول على المدرَّج التكراري لمجموعة بيانات التمرين /١/ من (١,٥,٥) دون استخدام جدول التوزيع التكراري وعلى أن نحدّد له عدد الفئات والتكرارات وتنفيذ بعض الإعدادات اللازمة (التي تحتاج إلى خبرة في عمل هذه البرنامج حيث لا مجال هنا لشرحها)، فنحصل على العرض الآتي للمدرَّج التكراري:




**Class Boundaries**

الشكل (٢٦, ١.د). المدرج التكراري لبيانات الجدول (١٧, ١)

٢- لتكن لدينا مجموعة البيانات الآتية التي تمثّل الطول لـ 68 نبتة قمح في مختبر زراعي (مقدّرة بالمليمتر) وذلك بعد شهرين

من عمليَّة الإنتاش:

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 168 | 159 | 172 | 171 | 189 | 180 | 152 | 195 | 185 | 165 | 167 | 170 | 155 | 173 | 194 |
| 188 | 185 | 172 | 199 | 152 | 186 | 170 | 167 | 208 | 167 | 153 | 168 | 179 | 177 | 171 |
| 186 | 167 | 193 | 165 | 150 | 172 | 195 | 140 | 173 | 169 | 180 | 187 | 184 | 175 | 167 |
| 181 | 184 | 170 | 160 | 169 | 155 | 166 | 190 | 200 | 173 | 203 | 160 | 145 | 171 | 177 |
| 181 | 191 | 182 | 172 | 196 | 179 | 154 | 148 | | | | | | | |

ولنقم بصبّ هذه البيانات في جدول توزيع تكراري فيه عدد الفئات $k = 7$، فنجد أنَّ السعة لكل فئة فعليّة من فئاته تساوي:

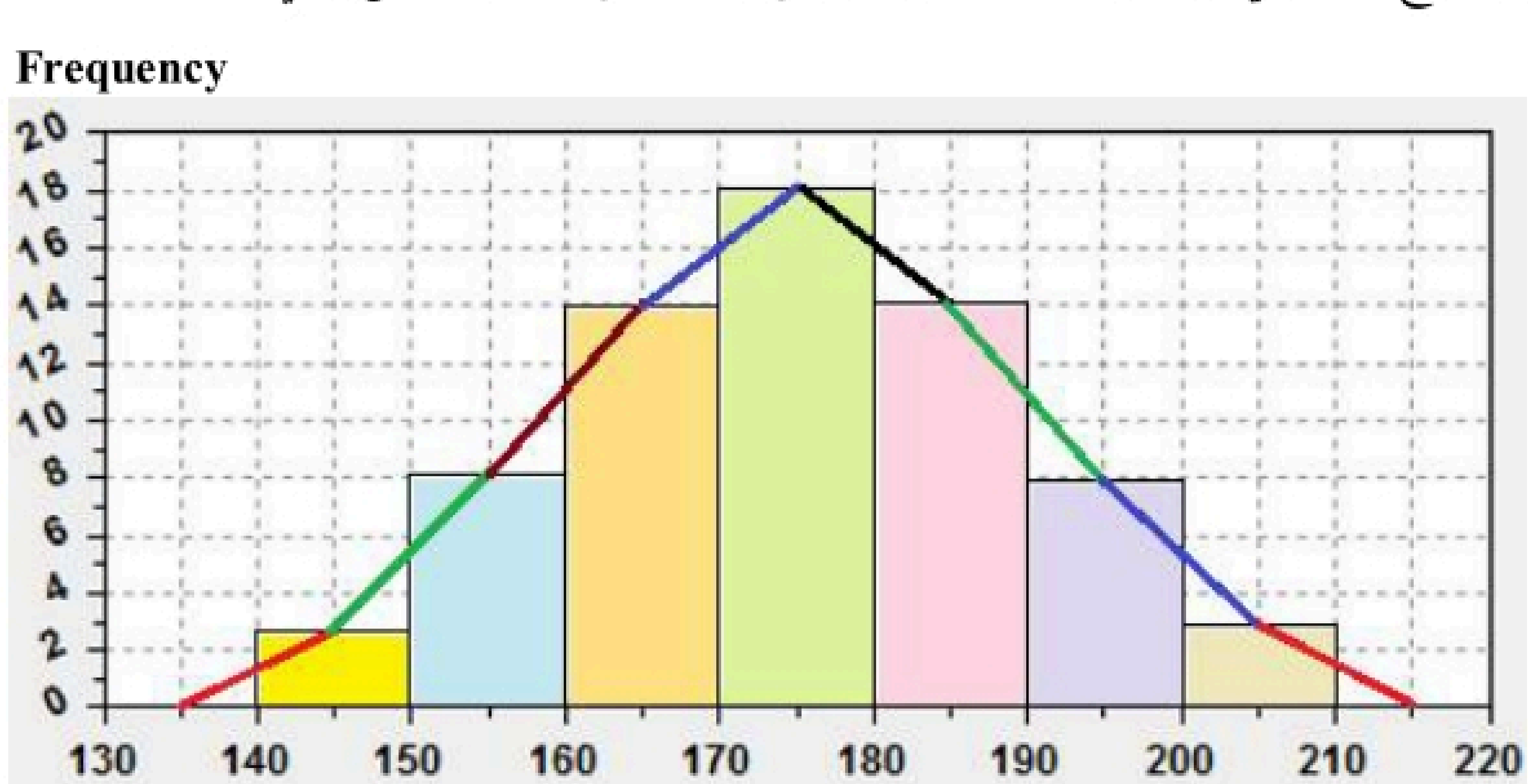
$$\mathbf{C} = (\mathcal{R} + 1) \,/\, k = 69 \,/\, 7 = 9.86$$

ولجعل عملية الصبّ أكثر سهولة ويُسراً سنقرّب قيمة السعة إلى العدد 10، فيكون لجدول التوزيع التكراري المطلوب العرض الآتي:

الجدول (١٨ , ١)

| رقم الفئة | الحدود الفعليّة للفئة | مركز الفئة | تكرار الفئة | التكرار المتجمّع الهابط للفئة |
|---|---|---|---|---|
| 1 | 140 → 150 | 145 | 3 | 3 |
| 2 | 150 → 160 | 155 | 8 | 11 |
| 3 | 160 → 170 | 165 | 14 | 25 |
| 4 | 170 → 180 | 175 | 18 | 43 |
| 5 | 180 → 190 | 185 | 14 | 57 |
| 6 | 190 → 200 | 195 | 8 | 65 |
| 7 | 200 → 210 | 205 | 3 | 685 |
| Total | ------------- | ---- | 68 | ------------- |

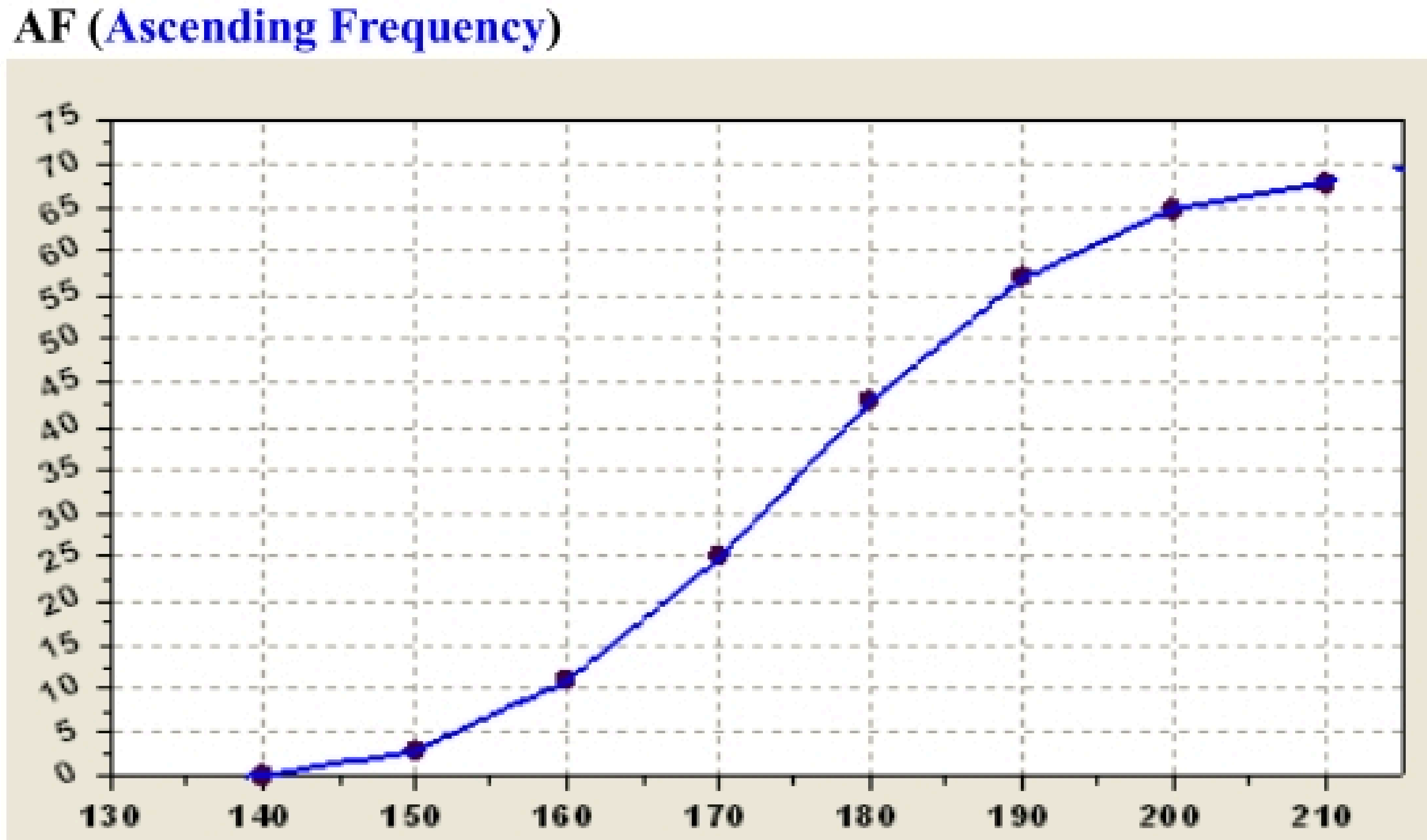
فعندئذٍ نجد للمدرَّج والمضلَّع التكراري لبيانات هذا الجدول العرض المقدّمة من خلال الشكل الآتي:




الشكل (٢٧ , ١ .أ)

وأخيراً يكون لمضلَّع التكرار المتجمّع الصّاعد الشكل الآتي:

AF (Ascending Frequency)




الشكل (٢٧, ١.ب)

٣- لتكن لدينا مجموعة البيانات الآتية التي تُمثّل عدد جلسات علاج فيزيائي تلقاها 81 مريضاً حتى استكمال الاستشفاء.

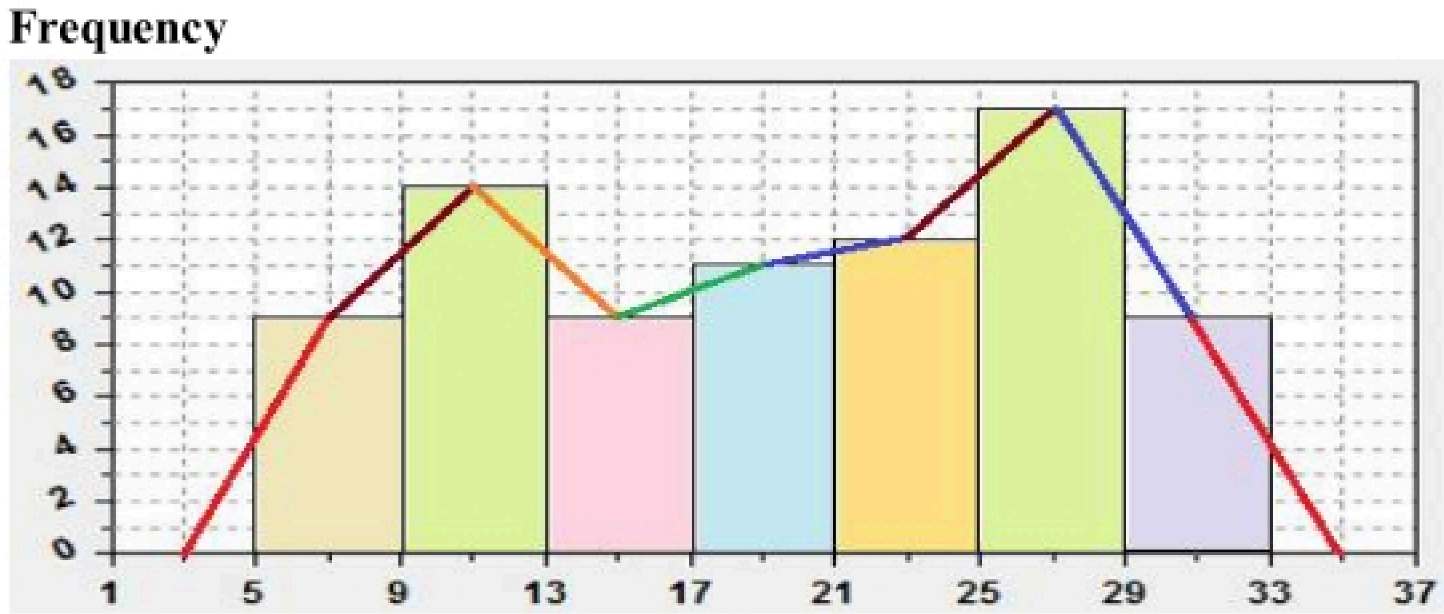
| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 11 | 24 | 6 | 9 | 30 | 7 | 13 | 26 | 20 | 30 | 20 | 11 | 14 | 6 | 9 | 29 | |
| 23 | 15 | 26 | 8 | 17 | 18 | 23 | 12 | 11 | 28 | 26 | 28 | 15 | 30 | 10 | 10 | |
| 22 | 32 | 5 | 16 | 25 | 7 | 22 | 20 | 12 | 13 | 25 | 24 | 5 | 28 | 20 | 27 | |
| 9 | 20 | 14 | 15 | 21 | 21 | 27 | 27 | 8 | 24 | 18 | 5 | 26 | 14 | 11 | 21 | |
| 22 | 10 | 32 | 27 | 27 | 17 | 20 | 24 | 18 | 29 | 30 | 26 | 31 | 26 | 10 | 11 | 25 |

فنجد أنَّ العدد الافتراضي للفئات $k = \lceil 6.34 \rceil = 7$ ، وأمّا المدى للبيانات فإنّه يساوي $\mathcal{R} = 27$ ، ومن ثمَّ تكون القيمة الافتراضية لسعة كل فئة هي $C = 4$ ، وبالتالي سيكون للجدول التكراري المطلوب العرض الآتي:

الجدول (١٩, ١)

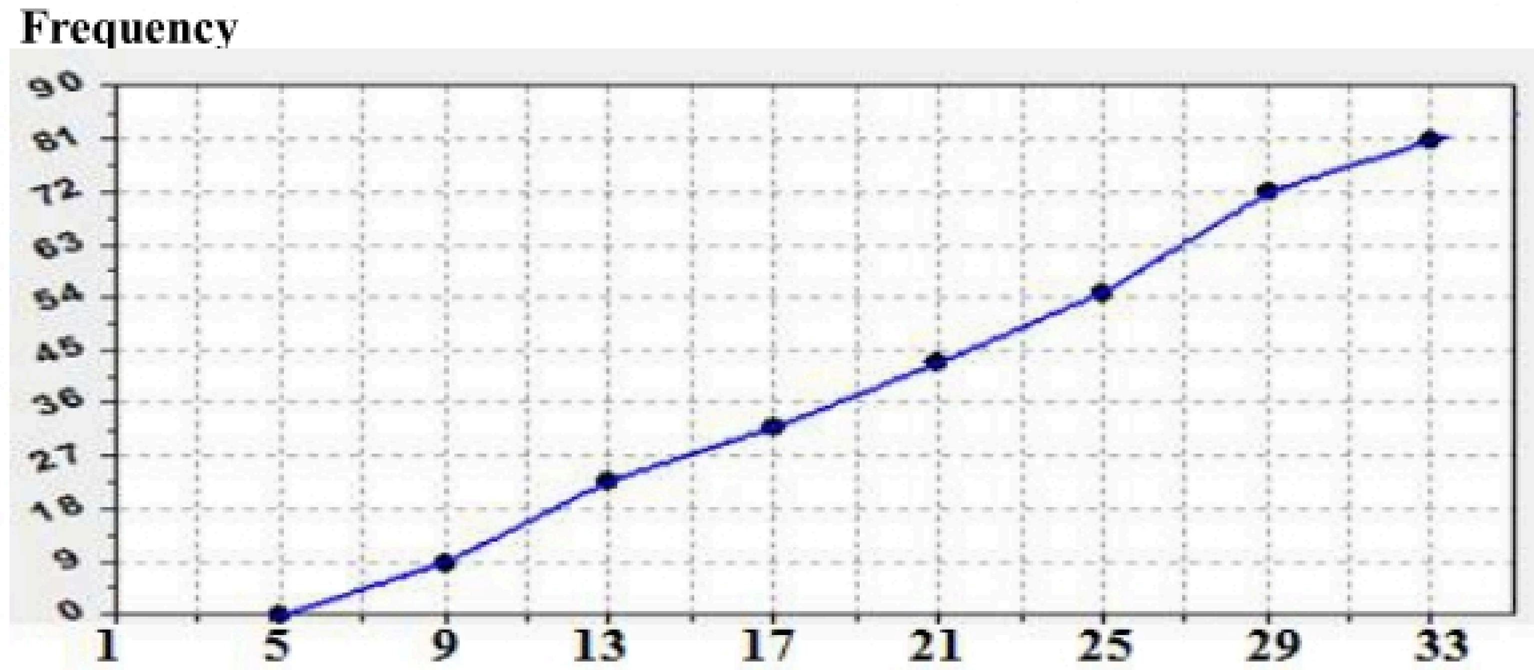
| رقم الفئة | الحدود الفعليّة للفئة | مركز الفئة | تكرار الفئة | التكرار المتجمّع الهابط للفئة |
|---|---|---|---|---|
| 1 | $05 \rightarrow 09$ | 7 | 9 | 9 |
| 2 | $09 \rightarrow 13$ | 11 | 14 | 23 |
| 3 | $13 \rightarrow 17$ | 15 | 9 | 32 |
| 4 | $17 \rightarrow 21$ | 19 | 11 | 43 |
| 5 | $21 \rightarrow 25$ | 23 | 12 | 55 |
| 6 | $25 \rightarrow 29$ | 27 | 17 | 72 |
| 7 | $29 \rightarrow 33$ | 31 | 9 | 81 |
| Total | ------------ | ---- | 68 | ------------ |

فعندئذ نجد للمدرَّج والمضلَّع التكراري لبيانات هذا الجدول العرض المقدّمة من خلال الشكل الآتي:




الشكل (٢٨ ,١ .أ)

وأخيراً يكون لمضلَّع التكرار المتجمّع الصّاعد الشكل الآتي:




شكل (٢٨ ,١ .ب)

## (١,٦) أشكال التوزيعات التكرارية

**Forms of Frequency Distributions**

لاحظنا من خلال الأمثلة السّابقة أنَّ التوزيعات التكرارية لها أشكال مختلفة ومتباينة، ولهذا وجَب تصنيف هذه التوزيعات لتسهيل التعامل مع مسمَّياتها وتصورها ذهنياً من مجرد معرفة نوعها، ولقد درجت العادة على تصنيف التوزيعات التكرارية من شكل التوزيع إلى نوعين رئيسيين:

### (١ ,٦ ,١) التوزيعات التكرارية المتناظرة Symmetric Frequency Distributions:

إنَّ الصفة التي تميّز هذه التوزيعات هي تناظر شكلها بالنسبة إلى محور مار من قمة التوزيع، وهذا النوع من التوزيعات يقسم بدوره إلى نموذجين هما:

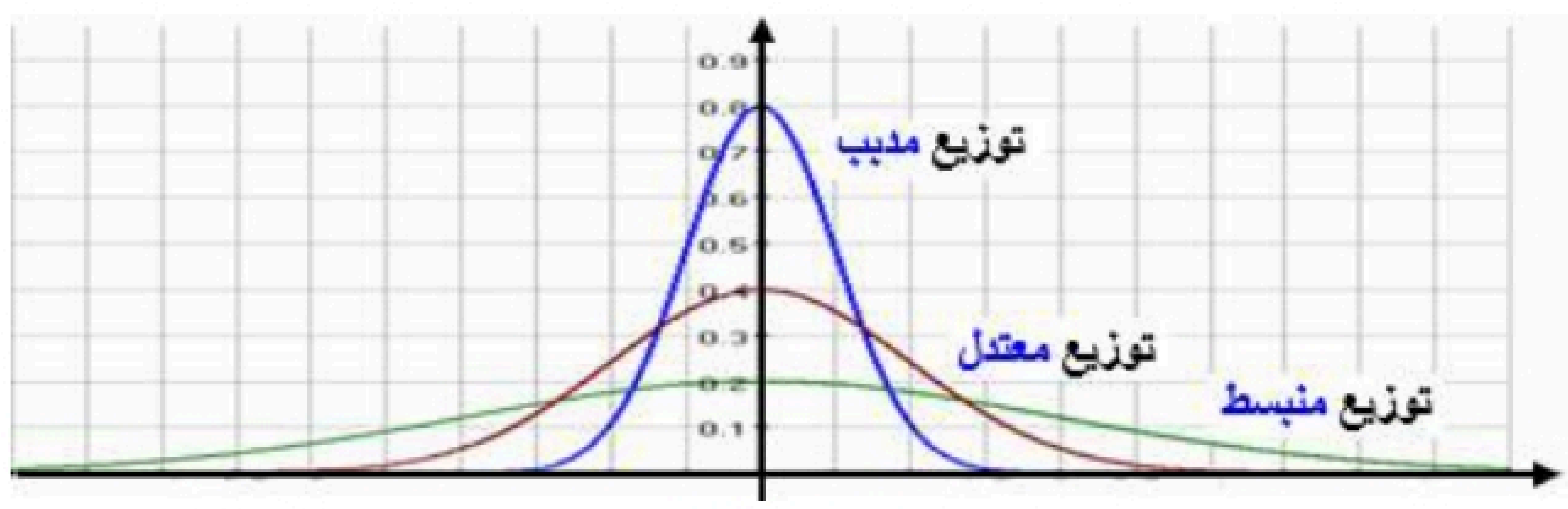
**١ – التوزيعات المتناظرة أحادية المنوال Unimodal Symmetric Distributions:**

التوزيعات المتناظرة الأحادية المنوال هي تلك التوزيعات التي يكون لشكلها البياني قمّة وحيدة، ومن أجل هذه التوزيعات يُمكننا التمييز بين ثلاثة أنواع هي:

أ – توزيعات مدببة Leptokurtic ب – توزيعات معتدلة Mesokurtic ج – توزيعات منبسطة Platykurtic.

علماً أنَّ مفهوم الاعتدال للتوزيعات بُنِيَّ على مدى تقارب شكله من شكل دالَّة الكثافة للتوزيع الطبيعي المعياري، حيث لُمعامل

التفلطح (سنأتي على ذكره في نهاية الفصل الثاني ) للتوزيع الطبيعي المعياري قيمة تساوي الصفر، ولذلك اُتخِذَ شكل دالَّة الكثافة الاحتمالية للتوزيع الطبيعي المعياري على أنَّه شكل معتدل تُقاس عليه بقية التوزيعات (سنأتي على ذكر دالَّة الكثافة للتوزيع الطبيعي المعياري في الفصل السادس، والشكل الآتي للتوزيع المعتدل يمثِّل رسمها البياني).



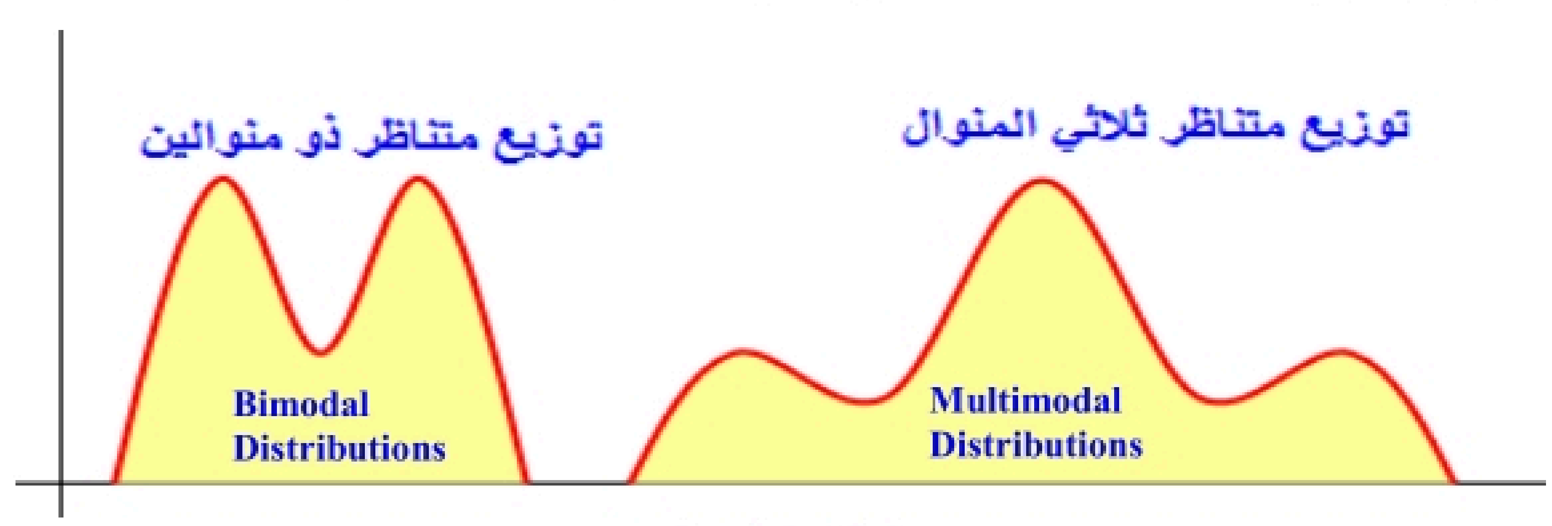

الشكل (١,٢٩).

وأمّا للحكم على تدبب أو تفلطح توزيعٍ ما فإنَّنا نحتاج لمعيار مُحدّد سنأتي على ذكره لدى تقديم المقاييس العددية لأشكال التوزيعات في الفصل القادم.

لاحظ أنَّ مجموعة بيانات التمرين / ٢/ من (١,٥,٥) السّابقة تقدّم لنا مثالاً على هذا النوع من التوزيعات، حيث يُلاحظ التناظر لشكل المدرج التكراري وكذلك لشكل المضلّع التكراري لتلك البيانات.

**٢ – التوزيعات المتناظرة متعدّدة المناويل Multimodal Symmetric Distributions**

تتميَّز هذه الأنواع من التوزيعات بأنَّها تمتلك قمتين على الأقل، والنموذجان الآتيان يوضِّحان لنا بعضاً منها:




الشكل (١,٣٠).

إنَّ مجموعة بيانات التمرين / ٣/ من (١,٥,٥) السّابقة تقدّم لنا مثالاً على توزيع تكراري ثنائي المنوال حيث يُلاحظ وجود قمتين لشكل المدرَّج التكراري وكذلك لشكل المضلّع التكراري لتلك البيانات، ولكنَّه ليس متناظراً.

### (١,٦,٢) التوزيعات التكرارية الملتوية Skewed Frequency Distributions

وهذه التوزيعات التكرارية لا تُحقِّق خاصية التناظر بالنسبة إلى محور مار من قمّة التوزيع في حال التوزيعات الأحادية المنوال، وفي حال التوزيعات المتعددة المناويل يكون شكل التوزيع غير متناظر بالنسبة إلى محور مار من منتصف قاعدة التوزيع (أي إنَّها توزيعات غير متناظرة Non Symmetric Distributions)، ويقسم هذا النوع من التوزيعات إلى صنفين رئيسيين هما:

**١ – التوزيعات الملتوية نحو اليمين (أو موجبة الالتواء) Skewed to the Right (or Positive Skewed) Distributions**

إنَّ أشكال هذه التوزيعات تتميَّز بظهور ذيل طويل مُمتد نحو اليمين أكثر من ذلك الذي يمتد نحو اليسار، والشكل الآتي يوضِّح ذلك من أجل توزيع وحيد المنوال.


الشكل (١,٣١)

٢ – **التوزيعات الملتوية نحو اليسار** (أو سالبة الالتواء) **Skewed to the Left (or Negative Skewed) Distributions**

إنَّ أشكال هذه التوزيعات تتميَّز بظهور ذيل طويل ممتد نحو اليسار أكثر من ذلك الذي يمتد نحو اليمين، والشكل الآتي يوضِّح ذلك من أجل توزيع وحيد المنوال.


الشكل (١,٣٢).

إنَّ مجموعة بيانات التمرين / ١/ من (١,٥,٥) السّابقة تقدّم لنا مثالاً على بيانات لها توزيع ملتوٍ نحو اليسار، حيث يُلاحظ الالتواء لشكل المدرج التكراري وكذلك لشكل المضلّع التكراري لتلك البيانات.

إنَّ كيفية الحكم على التواء التوزيع يحتاج إلى معايير مُحدَّدة سنأتي على ذكرها في الفصل القادم.

**(١,٦,٣) تفسير شكل التوزيع:**

في الواقع يمكن تقديم بعض التفسيرات لأشكال التوزيعات، ومنها:

١ – إذا كان للتوزيع التكراري شكل متناظر معتدل فإنَّ ذلك يعني أنَّ البيانات تتوزع طبيعياً على وجه التقريب، وأنَّ الأخطاء المرتكبة لدى عمليَّة القياس هي على الغالب أخطاء عشوائية (غير متعمَّدة).

٢ – إذا كان للتوزيع التكراري عدَّة مناويل فإنَّ ذلك يدلّل على وجود عدَّة أسباب فاعلة ومؤثِّرة في التجربة (أو المسألة) المولِّدة للبيانات، وعادة يكون عدد هذه الأسباب الفاعلة والمؤثِّرة مساوياً لعدد المناويل في التوزيع.

٣– إذا كان للتوزيع التكراري شكل ملتو نحو اليمين فإنَّ ذلك يعني على الغالب تنفيذ عمليات فرز بشأن البيانات بحيث يُستثنى منها البيانات التي تقلّ عن مقدار مُحدَّد تفّرضه طبيعة الدراسة الإحصائية، وأمّا إذا كان شكل التوزيع ملتو نحو اليسار فإنَّ ذلك يعني على الغالب تنفيذ عمليات فرز بشأن البيانات بحيث يُستثنى منها البيانات التي تزيد عن مقدار محدَّد تفرضُه طبيعة الدراسة الإحصائية.

٤ – إذا كان شكل التوزيع التكراري منبسطاً فإنَّ ذلك يعني على الغالب تنفيذ عمليات فرز بشأن البيانات بحيث يُستثنى منها البيانات التي تقل وتزيد عن قيمة محدَّدة تفرضها طبيعة الدراسة الإحصائية، وكذلك يدلِّل على أنَّ البيانات تتبعثر بعيداً عن قيمة المتوسط لهذه البيانات وذلك تبعاً لمقدار تفرطح هذا التوزيع.

**هذا ما تيسَّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل**

# تمارين الفصل الأول

١ - صنِّف المتغيِّرات الآتية وفقاً لكافة المعايير التي درستها في هذا الفصل مع تعيين مجموعة قيمها إن أمكن ذلك (يمكنك استخدام افتراضات مناسبة عند الضرورة):

أ - درجة الطالب في الاختبار الأخير لمقَّرر دراسي.

ب- الحرارة (مقدرة بالدرجة المئوية Celsius) داخل ثلاجة منزلية.

ج- أعمار الطلاب (مقدراً بالسنوات) في جامعة ما.

د- الزمن اللازم للوصول إلى العمل إذا كنت قد انطلقت من المنزل.

هـ- عدد الكواكب السيَّارة التي تدور حول النجوم في مجرَّة ما.

و- أطوال الطلاب في جامعة ما.

ز- عدد المقرَّرات الدراسية لطالب يدرس في جامعة ما.

ح- أوزان الماشية في مزرعة لتسمين الأغنام.

ط- أرقام البطاقات الشخصيّة للمواطنين في بلد ما.

ي- عدد مرَّات قذف قطعة نقود معدنية حتى الحصول على صورة لأول مرَّة.

ك- عدد الحوادث المرورية التي تقع على طريق سريع خلال شهر مُعيَّن من السنة.

ل- أعمار (مقدرة بالساعة) الأطفال حديثي الولادة (الخدَّج) في مستشفى للتوليد.

م - الحرارة (مقدرة بالكلفن Kelvin) داخل جهاز تبريد ذي قدرة عالية جداً على التبريد (يمكنه الوصول إلى الصفر المطلق ولو نظرياً).

ن- الحالات الجنائية في المحاكم.

ص- الألوان الناتجة عن تحليل الضوء الأبيض.

٢ - صنِّف البيانات الناتجة عن المتغيِّرات في السؤال السّابق وفقاً للمقاييس التي درستها في هذا الفصل مع مناقشة حالة الصفر ودلالته إن وجد.

٣- البيانات الآتية (كل قيمة مكوَّنة من ثلاثة أرقام) تمثِّل قوة الضغط (مقدَّرة بالرطل على البوصة) لعيِّنة مكوَّنة من78سبيكة معدنية لاستخدامها في تصنيع الدوائر المتكاملة.

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 123 | 134 | 178 | 76 | 167 | 184 | 105 | 229 | 146 | 218 | 157 | 101 |
| 171 | 105 | 221 | 183 | 186 | 121 | 181 | 180 | 143 | 97 | 154 | 153 |
| 174 | 120 | 104 | 115 | 160 | 208 | 158 | 133 | 207 | 180 | 190 | 93 |
| 194 | 133 | 156 | 168 | 167 | 141 | 245 | 208 | 74 | 199 | 181 | 158 |

والمطلوب صب هذه البيانات في جدول الساق والأوراق.

٤ - الجدول الآتي يعرض لنا قيمة المبيعات في مؤسسة خلال الأعوام 2005–2012 مقدّرة بـ مليون ريـال سعودي، والمطلوب تقديم العروض البيانية المختلفة (الأعمدة – الخطوط المنكسرة – الدائرة) لبيانات هذا الجدول.

| Year | 2012 | 2011 | 2010 | 2009 | 2008 | 2007 | 2006 | 2005 | العام |
|---|---|---|---|---|---|---|---|---|---|
| Amount | 180 | 145 | 250 | 210 | 175 | 100 | 200 | 150 | المبلغ |

٥- سُئل طلاب (عددهم 30) في كلية إدارة الأعمال عن تخصصاتهم، فكانت ردودهم كما يأتي:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **A** | **M** | **M** | **A** | **M** | **M** | **E** | **M** | **O** | **A** |
| **E** | **E** | **M** | **A** | **O** | **E** | **M** | **A** | **M** | **A** |
| **M** | **A** | **O** | **A** | **M** | **E** | **E** | **M** | **A** | **M** |

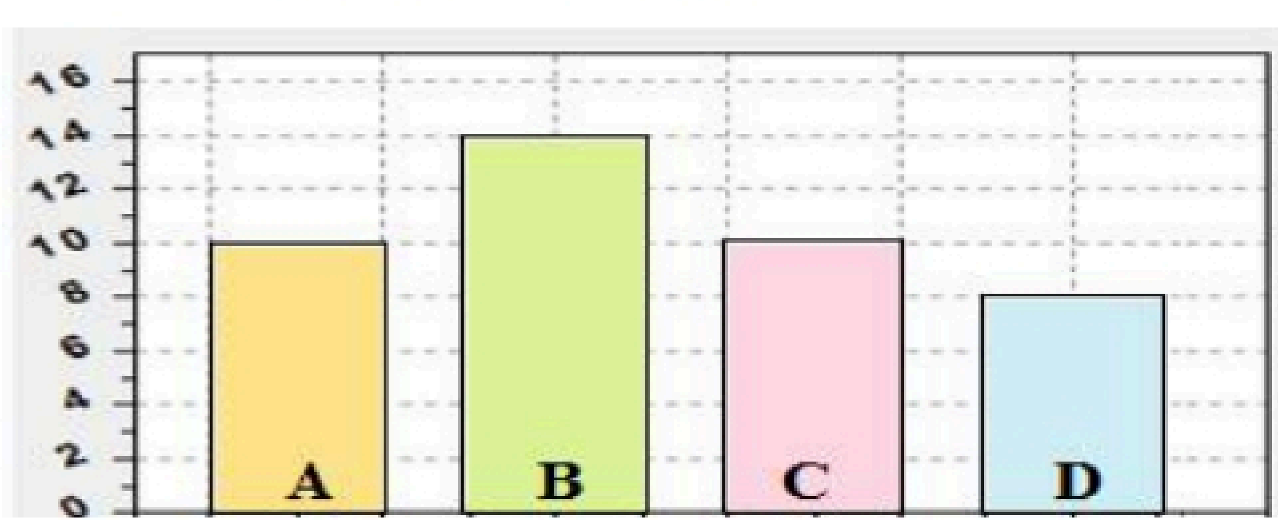
علماً أنَّ M تعني إدارة، وA تعني المحاسبة، وE تعني اقتصادًا، وأخيراً O تعني اختصاصاً آخر. عندئذ المطلوب بناء الجدول التكراري ومن ثمَّ تقديم العرض الشرائطي والدائري لهذه البيانات.

٦- طلب من 20 شخصاً من نزلاء فندق أن يدونوا مستوى الخدمة المقدَّمة لهم في هذا الفندق وذلك بكتابة إحدى العبارات الآتية: ممتازة Excellent، جيدة Good، متوسطة Average، ما دون المتوسطة Below Average أو ضعيفة Poor، فكانت نتائج إجاباتهم كما في الجدول الآتي:

| | | | | |
|---|---|---|---|---|
| **Below Average** | **Good** | **Average** | **Average** | **Poor** |
| **Good** | **Average** | **Average** | **Average** | **Good** |
| **Average** | **Good** | **Good** | **Average** | **Below Average** |
| **Good** | **Poor** | **Excellent** | **Excellent** | **Poor** |

والمطلوب بناء الجدول التكراري لهذه البيانات، ومن ثمَّ تقديم العروض البيانية المختلفة (الأعمدة - الخطوط المنكسرة - الدائرة) الموافقة لهذه البيانات.

٧- لنفترض أنَّه لدينا مجموعة من البيانات (لتقديرات A، B، C وD) مقدَّمة من خلال العرض البياني الآتي (العرض باستخدام الأعمدة)، والمطلوب تقديم العروض البيـانية المختلفة الأخرى (الجدول - الخطوط المنكسرة - الدائرة) لبيانات هذا العرض.




٨- لتكن لدينا مجموعة من البيانات مقدَّمة من خلال الجدول التكراري الآتي:

| رقم الفئة | الحدود الفعليّة للفئة | مركز الفئة | تكرار الفئة | التكرار النسبي للفئة | التكرار المئوي للفئة | التكرار المتجمّع الصاعد الفئة | التكرار المتجمّع الهابط الفئة |
|---|---|---|---|---|---|---|---|
| 1 | 14 → 17 | | 3 | | | | 90 |
| 2 | 17 → 20 | | | 0.2111 | | | |
| 3 | 20 → 23 | | | | 13.33% | | |
| 4 | 23 → 26 | | | | | 59 | |
| 5 | 26 → 29 | | | | 16.67% | | |
| 6 | 29 → 32 | | 16 | 0.1777 | | | |
| Total | | | | | | | المجموع |

والمطلوب ما يلي:

أ - أكمل مُعطيات هذا الجدول.

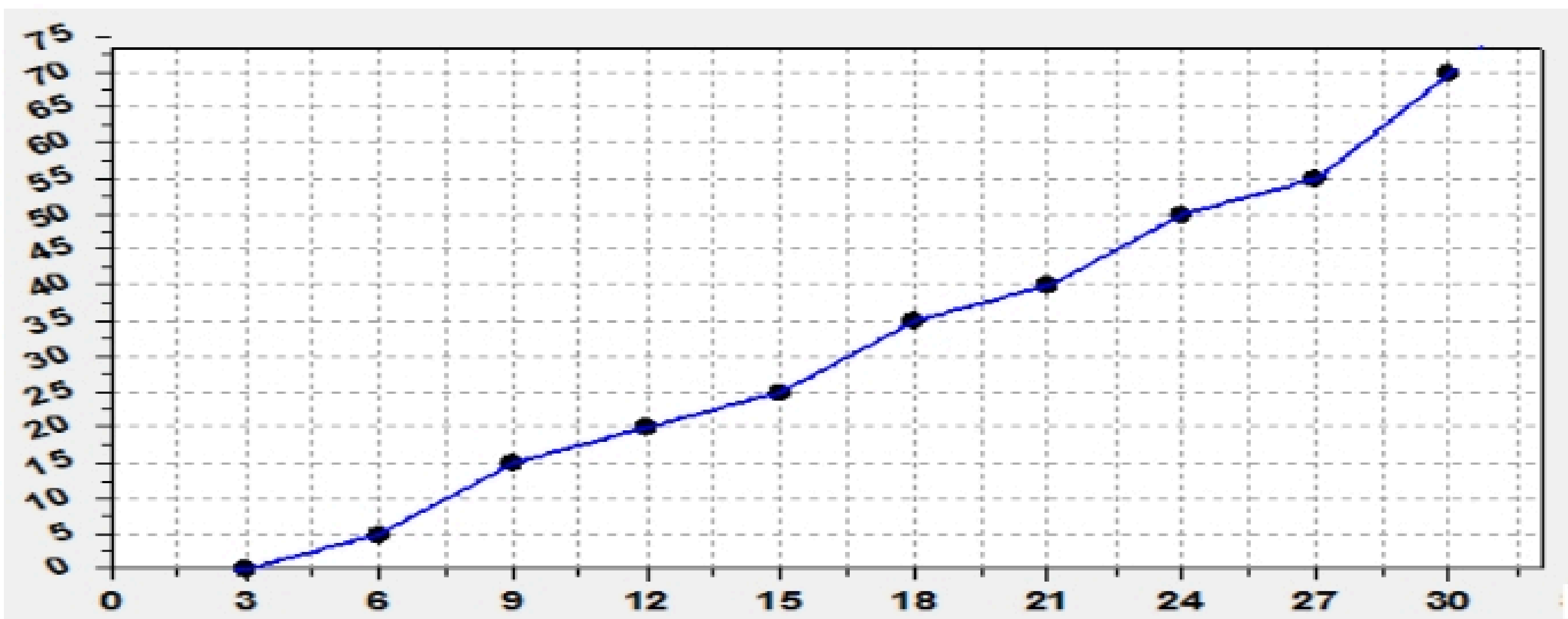
ب- قدِّم جميع العروض البيانية التي درستها في هذا الفصل لتمثيل بيانات هذا الجدول.

٩- لدينا مجموعة من البيانات مكوّنة من 180 مشاهدة، ومقدَّمة من خلال زوايا القطاعات الدائرية للقرص الدائري، حيث كانت قيم هذه الزوايا كما يلي:

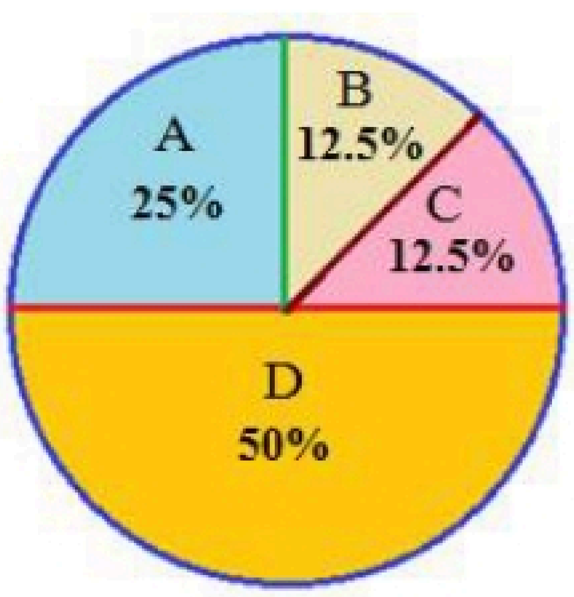
| الفئة | A | B | C | D | E |
|---|---|---|---|---|---|
| الزاوية المخصّصة لها | 36 | 54 | 90 | 162 | 18 |

والمطلوب رسم القرص الدائري للبيانات، ومن ثمَّ تقديم الجدول التكراري لهذه البيانات، وبعد ذلك رسم العرض الشرائطي والخطوط المنكسرة الموافقة لهذا العرض.

١٠- لدينا مجموعة من البيانات مقدَّمة من خلال مضلَّع التكرار المتجمِّع الصّـاعد الآتي:




والمطلوب تقديم العروض البيانية المختلفة الأخرى (جدول التوزيع التكراري، المدرَّج التكراري، المضلَّع التكراري ومضلَّع التكرار المتجمِّع الهابط) لبيانات هذا الجدول.




١١- لدينا مجموعة بيانات مقدَّمة من خلال القرص الدائري الآتي:

فإذا علمت أنَّ المجموع الكلي للبيانات يساوي 720 قيمة، فعندئذ المطلوب تقديم الجدول التكراري لهذه البيانات، ومن ثمَّ رسم العرض الشرائطي والخطّوط المنكسرة الموافقة لهذا العرض.

١٢- لتكن لدينا مجموعة من البيانات الخام مُعطاة كما يلي:

| | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 64 | 77 | 92 | 73 | 64 | 84 | 63 | 74 | 64 | 69 | 60 | 79 | 66 | 76 |
| 80 | 62 | 73 | 74 | 92 | 75 | 82 | 61 | 65 | 84 | 75 | 81 | 71 | 77 |
| 84 | 77 | 62 | 68 | 75 | 83 | 64 | 58 | 69 | 67 | 91 | 64 | 77 | 92 |
| 73 | 64 | 84 | 63 | 74 | 64 | 69 | 99 | 61 | 68 | 74 | 66 | 69 | 76 |

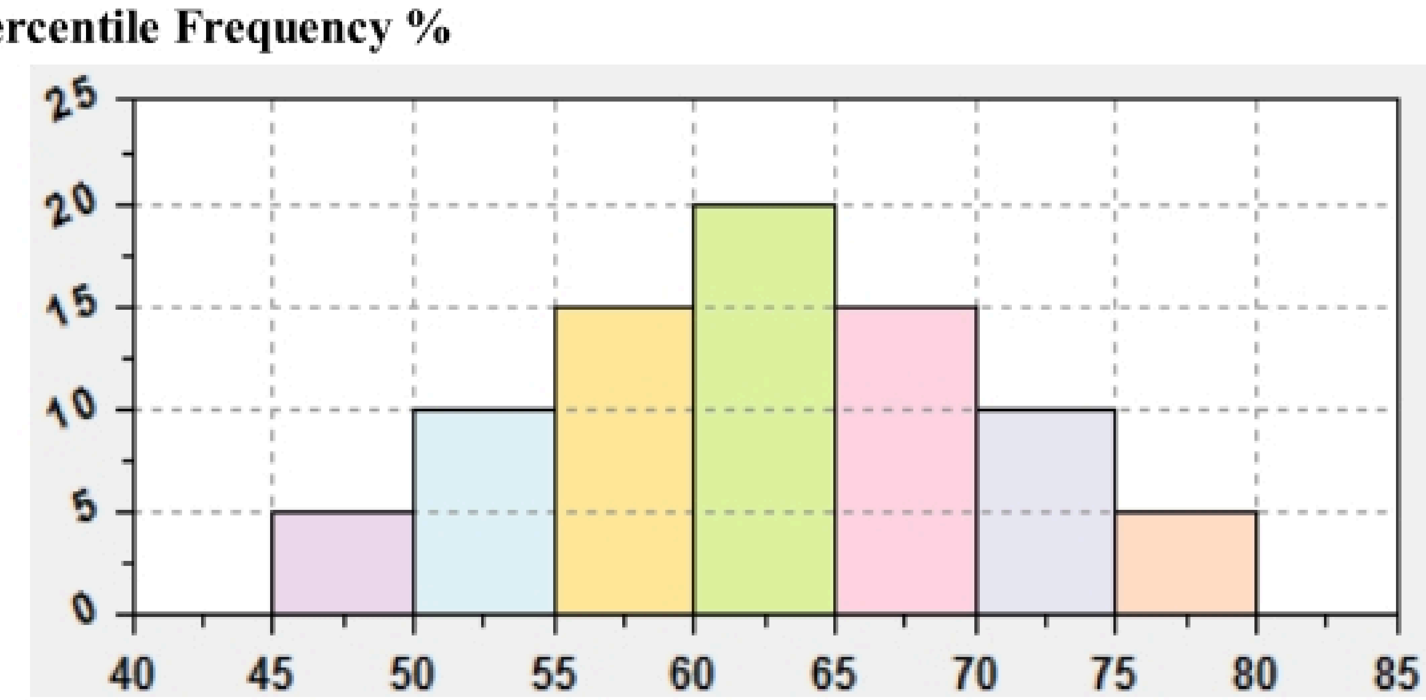
والمطلوب ما يلي:

أ - صبّ (أو تفريغ) هذه البيانات السابقة في جدول توزيع تكراري يحوي الأعمدة الأساسية فقط.

ب- ارسم المدرَّج والمضلَّع التكراري ومضلَّع التكرار المتَّجمع الصّاعد لبيانات جدول التوزيع التكراري النّاتج في الطلب السابق، ومن ثمَّ ناقش شكل التوزيع النّاتج.

١٣- لنفترض أنَّه لدينا مجموعة بيانات مقدَّمة من خلال المدرَّج التكراري الآتي:

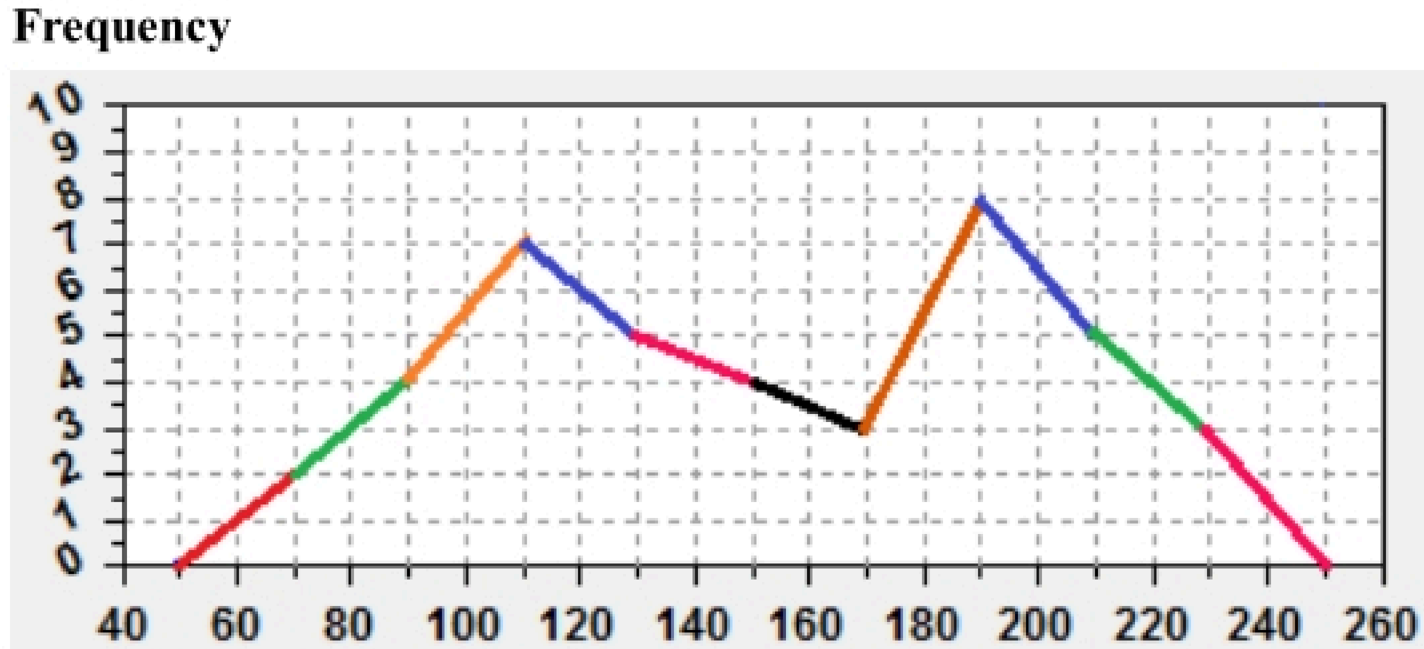
Percentile Frequency %




والمطلوب تقديم جدول التوزيع التكراري لهذه البيانات، ومن ثمَّ رسم المضلَّع التكراري، والمضلَّع التكراري، والتكراري النسبي، ومضلَّع التكرار المتجمّع الصّاعد لبيانات الجدول الناتج.

١٤- لنفترض أنَّه لدينا مجموعة بيانات مقدَّمة من خلال المضلَّع التكراري الآتي:

Frequency




والمطلوب تقديم جدول التوزيع التكراري لهذه البيانات، ومن ثمَّ رسم المدرَّج التكراري، والمدرَّج التكراري النسبي، والمدرَّج التكراري المئوي، ومضلَّع التكرار المتجمّع الصاعد والهابط لبيانات الجدول الناتج.

# الفصل الثاني

## المقاييس الوصفية للبيانات

## DESCRIPTIVE MEASURES OF DATA

### تمهيد

لقد ناقشنا في الفصل السّابق بعض الطرائق المستخدمة في عرض البيانات ولاحظنا أهميتها عند تقديم المعلومات، فهي تُعطي وصفاً عاماً وسريعاً للبيانات. إلاّ أنَّ فوائدها الاستقرائية قليلة، فالمضلَّع التّكراري لبيانات عيِّنة أُخذت من المجتمع يُعطينا تصوراً عن شكل المضلَّع التّكراري لمجتمع هذه البيانات، واستقراؤنا له يقف عند الفرض بأنَّه يوجد تشابه ما بين المضلَّع الممثِّل لبيانات هذه العيِّنة والمضلَّع الممثِّل لبيانات المجتمع الأصل، ومن ثمَّ المشكلة التي نقف عندها هي كيفية قياس درجة الاختلاف بينهما. إنَّ اعتماد المقارنة للأشكال فقط وباستخدام النظر قد لا يُعطينا الاستقراء الدقيق، وفي حال كان دقيقاً إلى حدٍّ ما فإنَّ ذلك يختلف من شخص إلى آخر حيث تصبح الدّقة هنا مفهومًا نسبيًّا. لهذا السبب يُفضَّل وجود مقاييس كميَّة تُمكننا من التّعرّف بشكل جيد على الخصائص والسلوك للتوزيع التّكراري الخاص بالمجتمع المدروس. لذلك سوف نوجّه اهتمامنا فيما يلي إلى تعيين قيمة أو أكثر من أجل مجموعة البيانات التي يُستدَلّ من خلالها على حقائق وخصائص الظاهرة التي تمثِّلها مجموعة البيانات ككلّ، وهذا يتطلب منا ما يلي:

**أولاً**: البحث عن قيمة عددية (أو قيم عددية) تَنْزع إليها البيانات (بمعنى أنَّ هذه القيمة- أو القيم- تبدو وكأنَّها نقطة جذب للبيانات)، وتُدعى مثل هذه القيم العددية **مقياساً للنزعة المركزيَّة** Central Tendency Measure، وفي حال وحدانية هذه القيمة من أجل مجموعة بيانات محدَّدة فإنَّه يُقال عنها إنَّها قيمة وسطى لهذه البيانات.

تجدر الإشارة هنا إلى أنَّ القيم العددية النّاتجة عن مقاييس النَّزعة المركزيَّة تندرج تحت مفهوم **مقاييس الموضع** للبيانات Position Measures of Data.

**ثانياً**: البحث عن قيمة عددية تُعبِّر عن مدى تشتّت (أو تبعثر) قيم البيانات حول مقياس نزعتها المركزّية (أو قيمتها الوسطى)، ويُدعى هذا العدد **مقياساً للتشتّت** Dispersion Measure.

**ثالثاً**: البحث عن قيمة عددية تُشير إلى شكل التوزيع للبيانات أو مظهره كما سبق وأوضحنا ذلك في الفصل الأول، وهنا نهتم بمعيّارين لشكل التوزيع هما:

**المعيّار الأول**: يتعلّق بتناظر شكل التوزيع من عدمه، أي الإجابة على السؤال الآتي:

**هل شكل التوزيع متناظر أم إنَّه ملتو (نحو اليمين أم نحو اليسار)؟**

إنَّ القيم العددية التي تُعبّر عن هذه الخصائص لشكلِ التوزيع تندرج تحت اسم **مقاييس الالتواء** Skewness Measures.

**المعيّار الثاني**: يتعلّق بارتفاع وانخفاض قمة شكل التوزيع، أي الإجابة على السؤال الآتي:

**هل التوزيع مدبباً، مفلطحاًأم معتدلاً ؟**

إنَّ القيم العددية التي تعبّر عن هذه الخصائص لشكل التوزيع تندرج تحت اسم **مقاييس التفلطح** (أو التفرطح) Kurtosis Measures.

يوجد في الواقع أنواع أخرى من المقاييس التي تساعدنا في عمليّة الاستقراء لنتائج البيانات، إلا أنَّها أقل أهميّة من المقاييس السّابقة. لذلك سنبدأ بتقديم شرح مفصَّل حول المقاييس التي نوهنا بها آنفاً، ومبتدئين بمقاييس النَّزعة المركزيَّة.

إنَّ مقاييس النَّزعة المركزيَّة تعدّ من أهمّ مقاييس الموضع، ولو أنَّ لمقاييس الموضع الأخرى (مثل الرُّبَيْعيّات **Quartiles** والمئينات **Percentiles**) عمل مهمّ أيضاً ولكنَّه مختلف بعض الشيء عن عمل مقاييس النَّزعة المركزيَّة، ولهذا السبب سنقدّم مقاييس النَّزعة المركزيَّة أولاً ومن ثمَّ الرُّبَيْعيّات والمئينات (وسوف نتجاوز مفهوم العُشَيرات **Deciles** بسبب قلّة استخدامه).

## (٢,١) مقاييس النزعة المركزية

### Central Tendency Measures

من المهام الأساسيّة في العديد من التحليلات الإحصائية عمليّة تقدير **مَعْلَمَة** Parameter الموضع لتمركز التوزيع (مَعْلَمَة تجمع على "مَعَالِم"، وتُعرف في بعض المراجع العربيّة باسم "وسيط" أيضاً، وتجمع على "وسطاء"). إنَّ عبارة **"النَّزعة المركزيَّة"** في إطار الدراسات الإحصائية تعني نزوع مجموعة قيم البيانات (أو مجموعة جزئية منها) نحو قيمة نموذجيّة أو مركزيّة تمثِّلها (أو تمثِّل هذه المجموعة الجزئية من البيانات)، وفي حال وحدانية هذه القيمة التي تنزع إليها البيانات فإنَّها تُماثل مفهوم مركز الثقل لجسم صلب في الميكانيكا، ولذلك فإنَّ عبارة النَّزعة المركزيَّة تتعلَّق بالطريقة التي تميل بها البيانات إلى التجمُّع حول قيمة ما (أو التكاثف حول مركز ثقلها).

#### (٢ , ١ , ١) تعريف (مقياس النَّزعة المركزيَّة)

إنَّ أي قيمة عدديّة (للبيانات الكميّة) أو رمز (للبيانات النوعيّة) تنزع إليها البيانات كلياً أو جزئياً تُدعى **مقياساً للنَّزعة المركزيَّة**.

بحسب هذا التعريف يتضِّح لنا أنَّه من أجل مجموعة بيانات مُعطاة قد يكون لدينا أكثر من قيمة ممثِّلة للنَّزعة المركزيَّة لهذه البيانات، ومفهوم المنوال الذي سيَرِد ذكره لاحقاً يوضِّح لنا معنى ذلك.

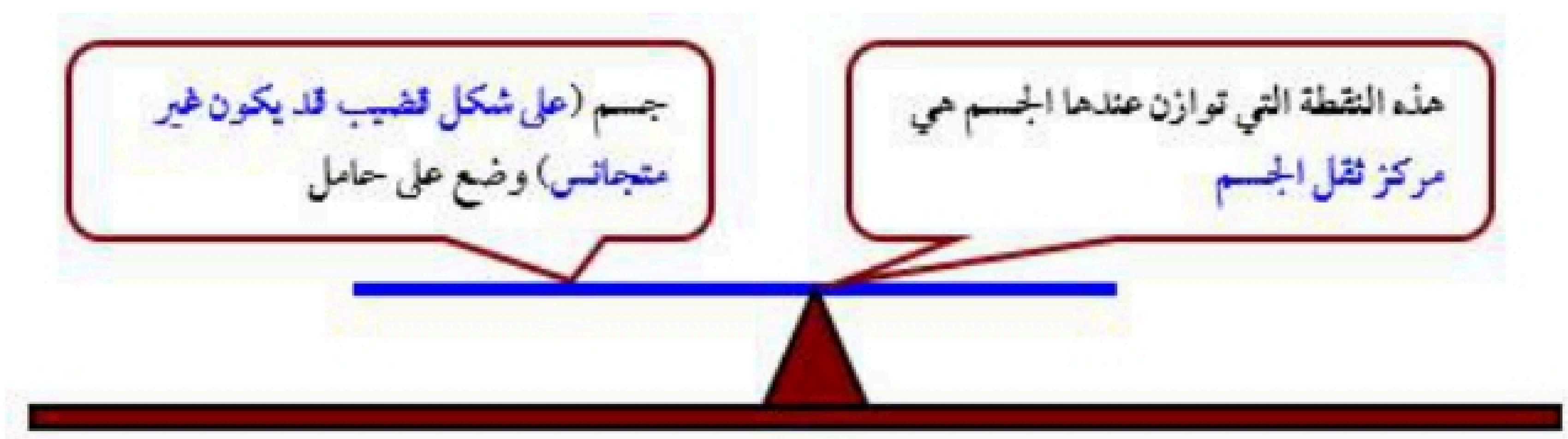
في عمليّة التحليل الإحصائيِّ غالبًا ما تُستخدم شروط قبل اختيار الشّكل الأوليّ للتحليل الإحصائي، ومنها على سبيل المثال: ماذا تعني كلمة القيمة النموذجية الممثِّلة للبيانات؟

لذلك فإنَّ الهدف الأوليّ قد يكون اختيار المقياس المناسب للنَّزعة المركزيَّة الذي يُعطينا دلالة واضحة لما تعنيه القيمة النموذجية الممثِّلة للبيانات. علماً أنَّ البيانات التي سنستخدمها في هذا الفصل والفصل الذي يليه هي بيانات وحيدة البُعد، أي إنَّها بيانات ناتجة عن متغيِّر وحيد البُعد، بمعنى أنَّه في إطار دراستنا هذه ستكون البيانات من الشكل $x_1$ ، $x_2$ ، ... و $x_n$ وليست على شكل ثنائيات $(x , y)$ أو ثلاثيات $(x , y , z)$ أو....

سنقدِّم فيما يلي العديد من مقاييس النَّزعة المركزيَّة، والمقياس الآتي يُعدّ من أشهره.

#### (٢ , ١ , ٢) المتوسط Mean (أو المتوسط الحسابي Arithmetic Mean)

إنَّ خير ما يوضِّح مفهوم المتوسط لمجموعة قيم عدديّة هو مركز الثقل لقضيب (ليس بالضرورة أن يكون متجانساً) يوضع على حامل. إنَّ مفهوم مركز الثقل لجملة مادية في الميكانيكا يُعرِّف مركز الثقل لجسم صلب على أنه تلك **النقطة** من الجسم التي يكون فيها مجموع العزوم بالنسبة لها يساوي الصفر وذلك عند وضع هذا الجسم في حقل قوى متوازي (انظر الرسم التوضيحي الآتي). فلو نظرنا إلى كل نقطة من هذا القضيب على أنَّها قيمة عددية (تعبِّر عن كتلة تلك النقطة)، فعندئذ مركز ثقل هذا القضيب (النقطة التي يتوازن عنها القضيب) هي قيمة المتوسط لهذه القيم. بالطبع لو كان القضيب متجانساً فعندئذ يكون مركز ثقله واقعاً على نقطة المنتصف للقضيب. لاحظ من هذا المفهوم يتضَّح لنا أن المجموع الجبري لانحرافات النقاط (القيم) عن مركز ثقلها (متوسطها) سيكون مساوياً للصفر.




الشكل (١ , ٢)

الآن، ومن أجل تعريف المتوسط (ويُعرف باسم "المتوسط الحسابي" Arithmetic mean أيضاً) الذي يُرمز له بـ $\overline{x}$ من أجل العينات وبـ $\mu$ من أجل المجتمعات، سنناقش الحالتين الآتيتين:

**أولاً: من أجل البيانات الخام Raw Data**، سنميِّز بين حالتين هما:

أ - إذا كانت $x_1$ ، $x_2$ ، ... و $x_n$ بيانات لعيِّنة، علماً أنَّ $n$ و N كيفي ولكن مُثبَّت، فعندئذ يحسب المتوسط لهذه البيانات يُحسب باستخدام العلاقة الآتية:

$$\overline{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i \qquad [2,1\text{-a}]$$

ب- إذا كانت $x_1$ ، $x_2$ ، ... و $x_N$ بيانات لمجتمع إحصائي، فإنَّ المتوسط لهذه البيانات يُحسب باستخدام العلاقة الآتية:

$$\mu = \frac{x_1 + x_2 + \dots + x_N}{N} = \frac{1}{N} \sum_{i=1}^{N} x_i \qquad [2,1\text{-b}]$$

**ثانياً: من أجل البيانات المُجمَّعة Grouped Data:**

بفرض أنَّه لدينا بيانات مجمَّعة مُعطاة في جدول توزيع التكراري (أو جدول تكراري حيث يكون لدينا ذات الجدول ولكن بدون عمود الفئات فقط، وبدلاً من مراكز الفئات يكون لدينا ممثلي البيانات بعد ترتيبها تصاعدياً ، ورقم الفئة هو رقم الممثِّل) كالآتي:

الجدول (١ , ٢)

| رقم الفئة | لحدود الفعليّة للفئة | مركز الفئة | تكرار الفئة | التكرار المتجمّع الصاعد للفئة |
|---|---|---|---|---|
| 1 | $b_0 \to b_1$ | $x_1$ | $f_1$ | $F_1 = f_1$ |
| 2 | $b_1 \to b_2$ | $x_2$ | $f_2$ | $F_2 = f_1 + f_2$ |
| ⋮ | ⋮ ⋮ ⋮ | ⋮ | ⋮ | ⋮ ⋮ ⋮ ⋮ |
| $k$-1 | $b_{k-2} \to b_{k-1}$ | $x_{k-1}$ | $f_{k-1}$ | $F_{k-1} = f_1 + f_2 + \dots + f_{k-1}$ |
| $k$ | $b_{k-1} \to b_k$ | $x_k$ | $f_k$ | $F_k = f_1 + f_2 + \dots + f_{k-1} + f_k$ |
| Total | ---------- | ----------- | $\sum f_i$ | ------ |

علماً أنَّنا كتبنا $\sum f_i$ بدلاً من $\sum_{i=1}^{k} f_i$ على سبيل التبسيط، فعندئذ إذا كانت البيانات لعيِّنة بمجموع تكرارات $\sum f_i = n$ (أو لمجتمع بمجموع تكرارات $\sum f_i = N$)، فإنَّ المتوسط لبيانات الجدول (سواءً كان جدول تكراري أو جدول توزيع تكراري) تُحسب كما يلي على الترتيب:

$$\overline{x} = \frac{1}{n} \sum_{i=1}^{k} f_i \cdot x_i \qquad \text{for sample} \qquad [2,2\text{-a}]$$

$$\mu = \frac{1}{N} \sum_{i=1}^{k} f_i \cdot x_i \qquad \text{for population} \qquad [2,2\text{-b}]$$

علماً أنَّ $x_i$ و $f_i$ هو مركز وتكرار الفئة ذات الرقم $i$ على الترتيب. نلاحظ هنا أنّه من أجل الجداول التكرارية ستكون قيمة المتوسط تساوي قيمة متوسط البيانات الخام نفسها دوماً، ولذلك سوف لن ندرج أمثلةً على حسابها من أجل هذا المقياس.

### (١ ,٢ ,١ ,٢) أمثلة

١ – لتكن لدينا البيانات الخام الآتية (132 مشاهدة) والممثِّلة لأطوال طلاب في مدرسة ثانوية (هنا المدرسة لها دور المجتمع) مقدَّرة بالسنتيمتر.

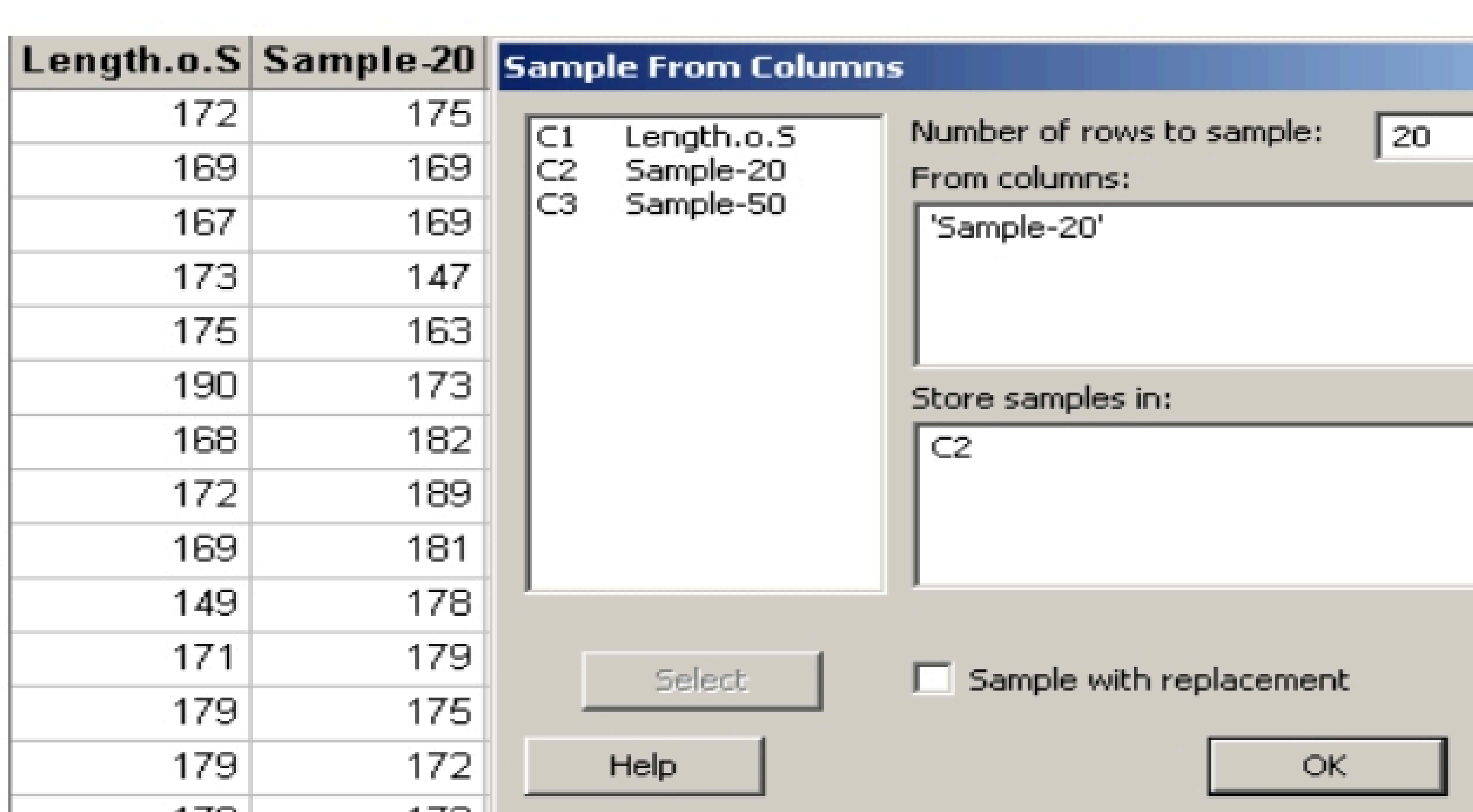
| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 172 | 167 | 178 | 196 | 188 | 154 | 171 | 181 | 188 | 169 | 149 | 195 |
| 167 | 157 | 169 | 175 | 193 | 189 | 167 | 192 | 178 | 182 | 160 | 175 |
| 172 | 173 | 173 | 173 | 179 | 193 | 157 | 152 | 173 | 160 | 172 | 168 |
| 175 | 168 | 179 | 173 | 183 | 169 | 195 | 181 | 178 | 190 | 158 | 190 |
| 175 | 169 | 172 | 173 | 160 | 173 | 168 | 172 | 172 | 158 | 163 | 169 |
| 170 | 178 | 172 | 172 | 170 | 183 | 189 | 167 | 184 | 173 | 181 | 177 |
| 169 | 165 | 185 | 175 | 173 | 180 | 193 | 163 | 172 | 149 | 172 | 178 |
| 173 | 155 | 175 | 175 | 173 | 163 | 171 | 172 | 173 | 171 | 160 | 163 |
| 173 | 182 | 147 | 179 | 160 | 177 | 175 | 180 | 179 | 160 | 192 | 169 |
| 179 | 173 | 171 | 171 | 178 | 175 | 173 | 169 | 171 | 170 | 189 | 185 |
| 170 | 169 | 168 | 158 | 171 | 179 | 152 | 147 | 180 | 173 | 163 | 171 |

عندئذ نجد أنَّ قيمة متوسط الطول لطلاب المدرسة يساوي:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{1}{132}\sum_{i=1}^{132} x_i = \frac{22812}{132} = 172.82$$

وبسحب عيِّنة عشوائية بسيطة بحجم $n = 20$ مستخدمين في ذلك برنامج Minitab نجد أنَّ لهذه العيِّنة العرض الآتي:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 175 | 169 | 169 | 147 | 163 | 173 | 182 | 189 | 181 | 178 |
| 172 | 173 | 177 | 173 | 181 | 178 | 193 | 173 | 179 | 175 |



ومن ثمَّ نجد قيمة المتوسط لبيانات هذه العيِّنة تساوي:

$$\overline{x}=\frac{1}{20}\sum_{i=1}^{20}x_i=\frac{3500}{20}=175$$

وبسحب عيِّنة عشوائية أخرى من هذه البيانات بحجم $n = 40$ مستخدمين في ذلك برنامج Minitab فنجد لها العرض الآتي:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 163 | 157 | 178 | 158 | 180 | 173 | 173 | 172 | 170 | 169 |
| 163 | 177 | 178 | 160 | 169 | 183 | 163 | 172 | 193 | 173 |
| 172 | 171 | 167 | 160 | 178 | 172 | 157 | 181 | 160 | 170 |
| 170 | 182 | 169 | 172 | 195 | 188 | 173 | 189 | 169 | 178 |

وأمّا قيمة المتوسط لبيانات هذه العيِّنة فنجدها تساوي:

$$\overline{x}=\frac{1}{40}\sum_{i=1}^{40}x_i=\frac{6897}{40}=172.43$$

فلاحظ هنا أنَّ قيمة المتوسط للعيّنات العشوائيّة ليس بالضرورة أن تتطابق مع قيمة المتوسط للمجتمع وإنَّما هي قيم تقديريّة لقيمة متوسط المجتمع، وسوف نرى فيما بعد (في الفصل التاسع ) أنَّه كلما كَبُر حجم العيِّنة أكثر فأكثر فإنَّ قيمة متوسطها تقترب أكثر فأكثر من قيمة متوسط المجتمع أيضاً.

٢- قِيسَ الضغطُ الانقباضي للقلب (في عمليّة قياس ضغط الدم مقدَّرة بـ mmHg ميليمتر زئبقي) لعيِّنة مكوَّنة من خمسين شخصاً مصاباً بارتفاعِ ضغط الدم، فكانت النتائج كما يلي:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 157 | 182 | 179 | 168 | 160 | 185 | 175 | 169 | 175 | 166 |
| 190 | 184 | 182 | 140 | 182 | 199 | 175 | 184 | 176 | 162 |
| 162 | 160 | 175 | 167 | 179 | 168 | 183 | 191 | 197 | 162 |
| 177 | 173 | 160 | 171 | 159 | 182 | 168 | 178 | 168 | 172 |
| 171 | 153 | 166 | 190 | 183 | 174 | 169 | 175 | 160 | 169 |

فنجد أنَّ قيمة المتوسط لهذه البيانات (الخام) تساوي:

$$\overline{x}=\frac{1}{50}\sum_{i=1}^{50}x_i=\frac{8652}{50}=173.04$$

ولنقم بتحميل هذه البيانات في جدول توزيع تكراري بحيث تكون سعة الفئة الفعليَّة تساوي 10.

**العمل**: من أجل ذلك نلاحظ أنَّ المدى لهذه البيانات يساوي:

$$\mathcal{R}=x_{\ell}-x_s=199-140=59$$

وبالتالي فأنَّ العدد الافتراضي للفئات يساوي $k=\lceil 5.6 \rceil = 6$ ، ومن ثَم يكون للبيانات الُمعطاة جدول التوزيع التّكراري الآتي:

الجدول (٢ , ٢)

| رقم الفئة | الحدود الفعليّة للفئة | مركز الفئة $x_i$ | تكرار الفئة $f_i$ | التكرار المتجمّع الصاعد للفئة $F_i$ | $x_i \cdot f_i$ |
|---|---|---|---|---|---|
| 1 | 140 → 150 | 145 | 1 | 1 | 145 |
| 2 | 150 → 160 | 155 | 3 | 4 | 465 |
| 3 | 160 → 170 | 165 | 17 | 21 | 2805 |
| 4 | 170 → 180 | 175 | 15 | 36 | 2625 |
| 5 | 180 → 190 | 185 | 9 | 45 | 1665 |
| 6 | 190 → 200 | 195 | 5 | 50 | 975 |
| Total | ------------- | ---- | $\sum f_i = 50$ | ------------------- | 8680 |

ومنه قيمة المتوسط لبيانات هذا الجدول تساوي:

$$\bar{x} = \frac{1}{50} \sum_{i=1}^{6} f_i \cdot x_i = \frac{8680}{50} = 173.60$$

إنَّ الفرق بين قيمتي المتوسط لكلٍ من البيانات الخام والمجمَّعة في جدول التوزيع التكراري ناتج عن عمليّة التجميع للبيانات في فئات، أي بسبب عدم التركيز على القيمِ المفردة للبيانات والاكتفاء بدور قيم مراكز الفئات كممثِّل لتجمعات جزئية من هذه البيانات (البيانات التي في الفئات) فقط.

### (٢ , ٢ , ١ , ٢) ملاحظات

١- إنَّ الكثير من صيغ العلاقات لها العرض نفسه من أجل بيانات العيِّنة والمجتمع، والخلاف بينهما في رمز حجم العيِّنة $n$ والمجتمع $N$ فقط. لذلك سنقدّم صيغ العلاقات من أجل بيانات العينات، ونذكر ما يُقابلها من رموز من أجل بيانات المجتمع مع ذكر الصيغة التفصيلية عند الضرورة (أي عندما يكون هناك خلاف في صيغة العلاقة) فقط.

٢- على الرغم من جودة المتوسط، وأنَّه من أهمّ وأكثر مقاييس النَّزعة المركزيَّة شيوعاً واستخداماً، إلاَّ أنَّه يُعاب عليه تأثُّره بالقيم المتطرفة Extreme Values (أو القيم المنعزلة Outlier Values وسنأتي على تعين هذه القيم لاحقاً بعد تقديم فقرة الرُّبَيْعيّات- انظر الملاحظة /٣/ من (٢ , ٢ , ٢ , ٢))، فعلى سبيل المثال لو قمنا بجمع تبرعات من فصل دراسي لصالح إحدى الجمعيات الخيريّة، وكانت لدينا البيانات الآتية مقدّرةً بالريـال السعودي:

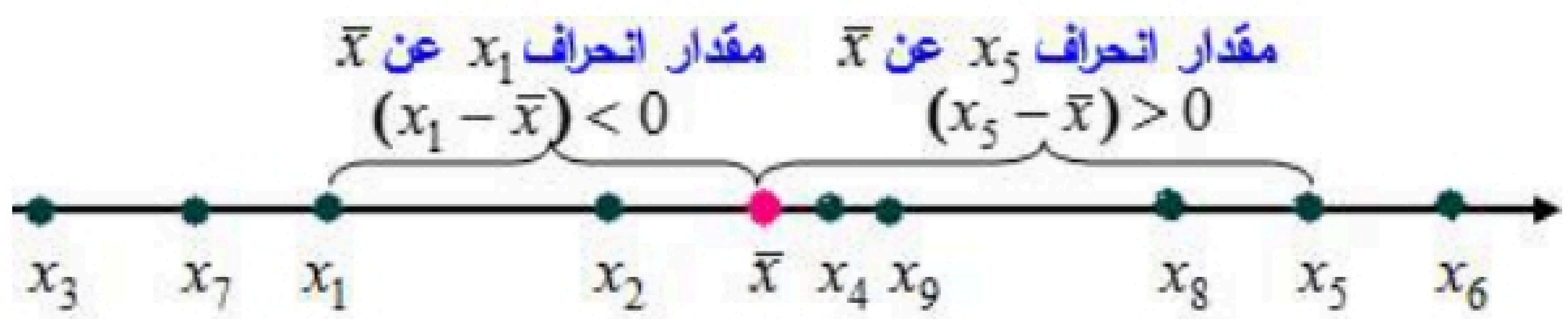
| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 7 | 4 | 7 | 6 | 5 | 4 | 3 | 5 | 100 | 4 | قيمة متطرفة |
| 4 | 3 | 7 | 6 | 5 | 5 | 7 | 4 | 5 | 5 | |
| 1 | 7 | 5 | 7 | 4 | 6 | 4 | 5 | 4 | 1 | |

فنجد أنَّ القيمة 100 متطرّفة (أو منعزلة)، وكذلك نجد أنَّ قيمة متوسط التبرعات لهذا الفصل يساوي 8 ر.س. رغم أنَّ جميع القيم أصغر من 8 باستثناء القيمة 100، فلو قمنا باستثناء القيمة المتطرفة (القيمة 100)، ومن ثمَّ حساب المتوسط لبقية القيم، لوجدنا أنَّها تساوي 4.83 ر. س، وهي تُعبِّر حقيقة عن القيمة الوسطى لتبرعات مُعظم التلاميذ، وهذا يعني أنَّ القيمة المتطرفة قد جذبت قيمة المتوسط باتجاهها فأعطانا تصوراً غير واقعي حول متوسط تبرعات التلاميذ.

٣- إنّ المجموع الجبري لانحرافات قيم بيانات خام $x_1$، $x_2$، ... و $x_n$ عن متوسطها $\overline{x}$ يساوي الصفر دوماً، وذلك لأنَّ:

$$\sum_{i=1}^{n}(x_i - \overline{x}) = \sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \overline{x} = \sum_{i=1}^{n} x_i - n \cdot \overline{x} = n \cdot \overline{x} - n \cdot \overline{x} = 0 \qquad [2,3]$$

انظر الرسم التوضيحي الآتي.




الشكل (٢,٢)

ولتوضيح ذلك حسابياً سنفترض أنّه لدينا البيانات 1 , 3 , 5 , 6, -5 , 6 , 3 , 9، فنجد أنَّ قيمة المتوسط لهذه البيانات هو $\overline{x} = 3.5$، ومن ثَمَّ يكون لدينا:

$$\sum_{i=1}^{n}(x_i - \overline{x}) = 5.5 - 0.5 + 2.5 + 2.5 - 8.5 + 1.5 - 0.5 - 2.5 = 0$$

## (٢,١,٣) المتوسط الموزون Weighted Mean

لتكن لدينا بيانات مُعطاة عددها $n$ بحيث إنَّ القيمة $x_i$ تكرَّرت $w_i$ مرَّة من أجل كل $i \in \{1,2,...,r\} =: \mathbb{N}_r$ (نذكّر هنا إلى أنَّ رموز مجموعات الأعداد المستخدمة في هذا الكتاب، ورموز أخرى مُدْرَجة، تجدها في بداية الملحق A) مع $\sum_{i=1}^{r} w_i = n$ ، فعندئذ المتوسط الموزون (أو المرجّح) لهذه البيانات يُرمز له بـ $\overline{x}_w$ من أجل العينات (وبـ $\mu_w$ من أجل المجتمعات)، يُعَّرف من خلال العلاقة الآتية:

$$\overline{x}_w = \frac{1}{\sum_{i=1}^{r} w_i} \sum_{i=1}^{r} w_i \cdot x_i = \frac{1}{n} \sum_{i=1}^{r} w_i \cdot x \qquad [2,4\text{-a}]$$

$$\mu_w = \frac{1}{\sum_{i=1}^{r} w_i} \sum_{i=1}^{r} w_i \cdot x_i = \frac{1}{N} \sum_{i=1}^{r} w_i \cdot x \qquad [2,4\text{-b}]$$

وفي هذه الحالة تُدعى الأعداد $w_1$، $w_2$، ... و $w_r$ أوزان البيانات $x_1$، $x_2$، ... و $x_r$ على الترتيب.

### (٢,١,٣,١) ملاحظة

نلاحظ من أجل البيانات المجمَّعة أنّ كل فئة من فئات جدول التوزيع التّكراري يمكن تمثيلها بمركزها، وأنَّ تكرار هذه الفئة يمثِّل الوزن لهذا المركز (وبالمثل بالنسبة إلى الجدول التّكراري حيث قيمة ممثِّل البيان لها دور مركز الفئة)، ومن ثَم فإنَّ قيمة المتوسط لبيانات جدول توزيع تكراري (جدول تّكراري) هي المتوسط الموزون لمراكز فئات (لممثِّلي بيانات) ذلك الجدول، فعلى سبيل المثال لو أخذنا البيانات الآتية:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 4.0 | 4.0 | 2.5 | 3.0 | 4.0 | 9.0 | 1.5 | 1.0 | 9.0 | 2.5 |
| 4.0 | 3.0 | 1.5 | 1.0 | 9.0 | 1.5 | 1.0 | 3.0 | 4.0 | 4.0 |
| 1.0 | 1.5 | 2.5 | 3.0 | 2.5 | 5.0 | 1.5 | 1.5 | 4.0 | 2.5 |

والتي يمكننا عرضها من خلال الجدول التّكراري (حيث الوزن $w_i$ يلعب دور التكرار هنا) الآتي:

الجدول (٢,٣)

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | sum |
|---|---|---|---|---|---|---|---|---|
| $x_i$ | 3 | 5 | 1 | 2.5 | 4 | 9 | 1.5 | ------ |
| $w_i$ | 4 | 1 | 4 | 5 | 7 | 3 | 6 | 30 |
| $x_i \cdot w_i$ | 12 | 5 | 4 | 12.5 | 28 | 27 | 9 | 97.5 |

فعندئذ نجد أنَّ قيمة المتوسط الموزون للبيانات المُعطاة هي:

$$\overline{x}_w = \frac{1}{\sum_{i=1}^{7} w_i} \sum_{i=1}^{7} w_i \cdot x_i = \frac{97.5}{30} = 3.25$$

٢- بفرض أنَّه لدينا:

$\overline{x}_1$ هو المتوسط لبيانات $x_{11}$، $x_{12}$، ... و $x_{1n_1}$

$\overline{x}_2$ هو المتوسط لبيانات $x_{21}$، $x_{22}$، ... و $x_{2n_2}$

..................................................

$\overline{x}_m$ هو المتوسط لبيانات $x_{m1}$، $x_{m2}$، ... و $x_{mn_m}$

فعندئذ تكون قيمة المتوسط لجميع البيانات السّابقة تساوي:

$$\overline{x} = \frac{n_1\overline{x}_1 + n_2\overline{x}_2 + \dots + n_m\overline{x}_m}{n_1 + n_2 + \dots + n_m} \qquad [2,5]$$

حيث نلاحظ أنَّه يساوي المتوسط الموزون لمتوسطات البيانـات المُعطاة بأوزان $n_1$، $n_2$، ... و $n_m$.

يقدّم لنا التعريف الآتي أحد مقاييس النَّزعة المركزيَّة المهمّة أيضاً (ويستخدم عادةً في مسائل النمو والسرعة).

### (٢, ١, ٤) المتوسط الهندسي Geometric Mean

من أجل تعريف المتوسط الهندسي لمجموعة بيانات سنناقش الحالتين الآتيتين، علماً أنَّنا سنرمز له بـ $\overline{x}_g$ من أجل العينات وبـ $\mu_g$ من أجل المجتمعات.

**١- من أجل البيانات الخام:**

لتكن لدينا $x_1$، $x_2$، ... و $x_n$ مجموعة بيانات موجبة تماماً (لعيِّنة أو لمجتمع حيث الفارق هنا رمز الدليل $n$ أو $N$ فقط)، فعندئذ يُعرَّف **المتوسط الهندسي** لهذه البيانات من خلال العلاقة الآتية:

$$\overline{x}_g = \sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} \qquad [2,6\text{-}a]$$

نلاحظ من الصيغة الأخيرة [2-6-a] أنَّه يُفضّل لحساب المتوسط الهندسي استخدام اللوغاريتمات، وذلك لأنَّه بأخذ لوغاريتم الطرفين للعلاقة الأخيرة نجد ما يلي:

أ- باستخدام اللوغاريتم الطبيعي نجد:

$$\ln \overline{x}_g = \frac{1}{n} \sum_{i=1}^{n} \ln x_i$$

ومن ثَمَّ ينتج لدينا العلاقة الآتية (والتي يُلاحظ فيها أنَّ الأس هو متوسط القيم $\ln x_1$، $\ln x_2$، ... و $\ln x_n$):

$$\bar{x}_g = \exp\left[\frac{1}{n}\sum_{i=1}^{n} \ln x_i\right] \qquad [2,6\text{-b}]$$

ب- باستخدام اللوغاريتم العشري نجد:

$$\log \bar{x}_g = \frac{1}{n}\sum_{i=1}^{n} \log x_i$$

فنحصل على الصيغة الآتية للمتوسط الهندسي:

$$\bar{x}_g = 10^{\frac{1}{n}\sum_{i=1}^{n}\log x_i} \qquad [2,6\text{-c}]$$

حيث نلاحظ أنَّ الأس هو متوسط القيم $\log x_1$، $\log x_2$، ... و $\log x_n$، وفي كلا الحالتين يتطلب انجاز الحسابات استخدام حاسبات أو جداول لوغاريتمية.

**٢- من أجل البيانات المجمَّعة:**

بفرض أنَّه لدينا بيانات مُجمَّعة في جدول توزيع تكراري (أو جدول تكراري) كما في الجدول (١ , ٢) وبمراكز فئات $x_1$، $x_2$، ... و $x_k > 0$ (أوبممثِّلين $x_1$، $x_2$، ... و $x_k > 0$)، فعندئذ يُحسب المتوسط الهندسي لبيانات ذلك الجدول من خلال العلاقة الآتية:

$$\bar{x}_g = \sqrt[\sum_{i=1}^{k} f_i]{x_1^{f_1}\cdot x_2^{f_2}\cdot \ldots \cdot x_k^{f_k}} \qquad [2,7\text{-a}]$$

علماً أنَّ $x_i$ و $f_i$ هو مركز وتكرار الفئة (الممثِّلين للبيانات وتكراراتها المقابلة) ذات الرقم $i$ على الترتيب، ولتبسيط العمليات الحسابية نأخذ لوغاريتم الطرفين للعلاقة الأخيرة فيكون:

$$\ln \bar{x}_g = \frac{1}{\sum_{i=1}^{k} f_i}\ln\left(\prod_{i=1}^{k} x_i^{f_i}\right) = \frac{1}{\sum_{i=1}^{k} f_i}\sum_{i=1}^{k} \ln x_i^{f_i} = \frac{1}{\sum_{i=1}^{k} f_i}\sum_{i=1}^{k} f_i \cdot \ln x_i$$

ومن ثَّم ينتج لدينا:

$$\bar{x}_g = \exp\left[\frac{1}{\sum_{i=1}^{k} f_i}\sum_{i=1}^{k} f_i \cdot \ln x_i\right] \qquad [2,7\text{-b}]$$

وهنا يُلاحظ أنَّ الأس هو المتوسط للقيم $f_1 \cdot \ln x_1$، $f_2 \cdot \ln x_2$، ... و $f_k \cdot \ln x_k$.

**(١ , ٤ , ١ , ٢) ملاحظة**

بفرض أنَّ $x_1$، $x_2$، ... و $x_r$ بيانات مُعطاة بأوزان $w_1$، $w_2$، ... و $w_r$ على الترتيب، فعندئذ يُعرَف المقدار الآتي:

$$\bar{x}_{g_w} = \sqrt[\sum_{i=1}^{r} w_i]{x_1^{w_1}\cdot x_2^{w_2}\cdot \ldots \cdot x_r^{w_r}} \qquad [2,8]$$

باسم **المتوسط الهندسي الموزون** Weighted Geometric Mean، ومن ثَّم يكون المتوسط الهندسي لبيانات جدول تكراري ذو $k$ فئة هو المتوسط الهندسي الموزون لمراكز فئات هذا الجدول على افتراض أنَّ التكرارات هي أوزان هذه المراكز.

تجدر الإشارة هنا إلى أنَّ المتوسط الهندسي يُستخدم في قطاعات رياضياتية تطبيقيّة عديدة منها التحويل والإدارة الماليّة والإحصاء السكاني.

**(٢, ١, ٤, ٢) أمثلة**

١- إذا كان مُعدَّل التضخّم في بلد ما هو 3 % سنة 2013، وكذلك 4.5 % سنة 2014، وأخيراً 7 % سنة 2015، فعندئذ يُحسب مُعدَّل التضخّم للسنوات الثلاث من خلال المتوسط الهندسي لهذه النسب، ويكون لدينا $\bar{x}_g = \sqrt[3]{3\times 4.5\times 7} = 4.55$، وهذا يعني أنَّ مُعدَّل التضخّم للسنوات الثلاث 4.55 %، وأمّا إذا أردنا حسابه باستخدام اللوغاريتمات فإنّنا سنحصل على النتيجة نفسها، حيث لدينا:

$$\bar{x}_g = \exp\left[\frac{1}{n}\sum_{i=1}^{n} \ln x_i\right] = \exp\left[\frac{1}{3}(\ln 3 + \ln 4.5 + \ln 7)\right]$$
$$= \exp\left[\frac{1}{3}(1.099 + 1.504 + 1.946)\right] = e^{1.516} = 4.55$$

٢- إذا كان مُعدَّل النمو السكاني لبلد ما هو 3.25 % في سنة 2011، و2.75 % في سنة 2012 و2.5 % في سنة 2013 وأخيراً 2.25 % في سنة 2015، فعندئذ يُحسب مُعدَّل النمو السكاني للسنوات الأربع من خلال المتوسط الهندسي، ونجد أنَّ:

$$\bar{x}_g = \sqrt[4]{3.25\times 2.75\times 2.5\times 2.25} = 2.66$$

أي إنَّ مُعدَّل النمو السكاني للسنوات الأربع يساوي 2.66%.

يقدّم لنا التعريف الآتي أحد مقاييس النَّزعة المركزيَّة الشهيرة أيضاً (**ويستخدم عادةً في مسائل التدفق**).

**(٢, ١, ٥) المتوسط التوافقي Harmonic Mean**

لتكن $x_1$، $x_2$، ... و $x_n$ بيانات خام مُعطاة وبحيث يكون $x_i \neq 0$ من أجل كل $i \in \mathbb{N}_n$، وكذلك $\sum_{i=1}^{n} x_i^{-1} \neq 0$، فعندئذ يُعرَّف **المتوسط التوافقي** لهذه لبيانات من خلال العلاقة الآتية:

$$\bar{x}_h = \frac{1}{\frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i}} = n\cdot\left(\sum_{i=1}^{n}\frac{1}{x_i}\right)^{-1} \qquad [2,9\text{-a}]$$

$$\mu_h = \frac{1}{\frac{1}{N}\sum_{i=1}^{N}\frac{1}{x_i}} = N\cdot\left(\sum_{i=1}^{N}\frac{1}{x_i}\right)^{-1} \qquad [2,9\text{-b}]$$

علماً أنّنا استخدمنا الرمز $\bar{x}_h$ من أجل العينات و $\mu_h$ من أجل المجتمعات.

**(٢, ١, ٥, ١) ملاحظات**

١- من العلاقة [2,9-a] نجد أنَّ:

$$\frac{1}{\bar{x}_h} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{x_i} \qquad [2,10]$$

وهذا يعني أنَّ مقلوب المتوسط التوافقي لمجموعة بيانات يساوي المتوسط لمقلوب البيانات نفسها.

٢- في الحالة الخاصّة عندما يكون لدينا $n=2$ و $n=3$ و $n=4$ فإنَّه يمكننا حساب المتوسط التوافقي $\bar{x}_h$ من خلال العلاقات الآتية على الترتيب:

$$\bar{x}_h = 2\,\frac{x_1 \cdot x_2}{x_1 + x_2}$$

$$\bar{x}_h = 3\,\frac{x_1 \cdot x_2 \cdot x_3}{x_1 \cdot x_2 + x_1 \cdot x_3 + x_2 \cdot x_3}$$

$$\bar{x}_h = 4\,\frac{x_1 \cdot x_2 \cdot x_3 \cdot x_4}{x_1 \cdot x_2 \cdot x_3 + x_1 \cdot x_2 \cdot x_4 + x_1 \cdot x_3 \cdot x_4 + x_2 \cdot x_3 \cdot x_4}$$

وهكذا دواليك.

٣- إنَّ العلاقة التي تربط بين المتوسطات $\bar{x}$ و $\bar{x}_g$ و $\bar{x}_h$ هي:

$$\bar{x}_h \le \bar{x}_g \le \bar{x} \qquad [2,11]$$

ومن أجل الحالة الخاصّة $n=2$ يكون لدينا:

$$\bar{x}_h = \frac{(\bar{x}_g)^2}{\bar{x}} \iff \bar{x}_g = \sqrt{\bar{x} \cdot \bar{x}_h} \qquad [2,12]$$

٤- تجدر الإشارة هنا إلى أنَّ المتوسط التوافقي يُستخدم عندما تتعلَّق قيم البيانات ببعض الوحدات، ومن الأمثلة الأكثر شيوعًا على ذلك حساب متوسط السرعة والتدفق، والمثالان الآتيان يوضّحان ذلك.

### (٢,٥,١,٢) أمثلة

١- تقطع سيارة مسافة مقدارها 10 كم على طريق خارجي بسرعة مقدارها 110 كم/ ساعة، وتتابع على طريق داخلي سريع لتقطع مسافة مقدارها 10 كم بسرعة 65 كم/ ساعة، وأخيراً تسير مسافة مقدارها 10 كم على طريق داخلي في حي شعبي بسرعة 30 كم/ ساعة، فعندئذ يكون متوسط السرعة لهذه السيارة على المسافة المقطوعة (**30 كم**) هو:

$$\bar{x}_h = 3\Big/\left(\frac{1}{110} + \frac{1}{65} + \frac{1}{30}\right) = 51.895\ km/h$$

٢- لدينا ثلاثة صنابير مياه تصبّ ماءً في حوض بحيث إنَّ الصنبور الأول والثاني والثالث يصب في الساعة 700، 900 و1200 لتر في الساعة على الترتيب، ولنحسب متوسط صبّ الصنابير الثلاثة بآن واحد خلال ساعة واحدة.

إنَّ متوسط صبّ الصنابير الثلاثة (مقدَّرةً باللیتر في الساعة) بآن واحد يُحسب باستخدام المتوسط التوافقي، فنجد أنَّ:

$$\bar{x}_h = 3\Big/\left(\frac{1}{700} + \frac{1}{900} + \frac{1}{1200}\right) = 889.412$$

ومنه يكون متوسط صبّ هذه الصنابير يساوي 889.412 لترًا في الساعة.

من المُلاحظ أنَّ السّمة المشتركة بين $\bar{x}$ و $\bar{x}_g$ و $\bar{x}_h$ هي اعتماد هذه المتوسطات على جميع قيم البيانات، بمعنى أنَّه إذا فقدت إحدى البيانات أو بعضها فإنَّها تصبح غير قابلة للتطبيق، ولهذا كان لا بدَّ من البحث عن مقياس بديل للنَّزعة المركزيَّة يساعدنا في تجاوز هذه العقبة.

يقدَّم لنا التعريف الآتي أحد مقاييس النَّزعة المركزيَّة، وهذا المقياس يُساعدنا في تجاوز مشكلة القيم المتطرفة وفقدان بعض البيانات (ولكن تحت شروط مُحَدَّدة)، أي إنَّه يقدّم لنا حلاً جزئياً لمشكلة فقدان بعض البيانات.

### (٦ , ١ , ٢) الوسيط Median

يُعرَّف **الوسيط** (وفي بعض المراجع يذكر باسم "الوسط") لمجموعة من البيانات على أنَّه القيمة التي يقع على يسارها 50% من البيانات وعلى يمنها 50% من البيانات أيضاً، وذلك بعد ترتيب البيانات تصاعدياً (أو تنازلياً)، علماً أنَّنا سنرمز للوسيط بـ $\tilde{x}$ من أجل العينات وبـ $\tilde{\mu}$ من أجل المجتمعات، ولحسابه سنناقش الحالتين الآتيتين:

**١ - من أجل البيانات الخام:**

نفرض أنَّ $x_1$ ، $x_2$ ، ... و $x_n$ بيانات خام مُرتّبة تصاعدياً (يمكن أن تكون مرتبة تنازلياً أيضاً). عندئذ يُحسب الوسيط لهذه البيانات باستخدام إحدى العلاقتين الآتيتين:

أ- إذا كان حجم البيانات (أو عدد البيانات) $n$ عدداً فردياً، فعندئذ يكون لدينا:

$$\tilde{x} := x_{\frac{n+1}{2}} \qquad [2,13\text{-a}]$$

ب- إذا كان حجم البيانات (أو عدد البيانات) $n$ عدداً زوجياً، فعندئذ يكون لدينا:

$$\tilde{x} := \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} \qquad [2,13\text{-b}]$$

**٢- في حالة البيانات المجمَّعة:**

بفرض أنَّه لدينا بيانات مقدَّمة من خلال جدول التوزيع التكراري (١ , ٢)، فعندئذ يُحسب **الوسيط** لبيانات ذلك الجدول من خلال العلاقة الآتية:

$$\tilde{x} := \tilde{L} + \frac{\frac{\sum f_i}{2} - \left(\tilde{F} - \tilde{f}\right)}{\tilde{f}} \times \mathbf{C} \qquad [2,14]$$

علماً أنَّ $\tilde{L}$ هو الحد الأدنى للفئة الفعليَّة الوسطيّة (علماً أنَّ الفئة الوسطيّة هي أول فئة تكرارها المتجمِّع الصاعد أكبر أو يساوي نصف مجموع التكرارات)، $\tilde{f}$ هو تكرار الفئة الوسطيّة، $\tilde{F}$ هو التكرار المتجمّع الصاعد للفئة الوسطيّة، و $\mathbf{C}$ هي سعة الفئة الوسطيّة (هنا لم نرفق دليلاً لـ $\mathbf{C}$ أو رمزاً خاصًا بها لأنَّنا نتعامل مع فئات متساوية السعة).

أمّا لو كانت البيانات مقدَّمة من خلال جدول تكراري، فعندئذ نعيد تنظيم الجدول بحيث تصبح قيم الممثِّلين في الجدول التكراري (١ , ٢) مرتّبةً تصاعدياً $x_1 < x_2 < \ldots < x_k$. بعد ذلك تُعيَّن قيمة الوسيط كما يلي:

أ- إذا كان مجموع التكرارات $\sum f_i = n$ فردياً فإنَّنا نأخذ الممثِّل $x_i$ الموافق لأصغر تكرار تراكمي أكبر أو يساوي $\frac{n+1}{2}$.

ب- إذا كان مجموع التكرارات $\sum f_i = n$ زوجياً، وكان لممثِّل $x_i$ تكرار تراكمي يساوي $\frac{n}{2}$ تماماً، فإنَّ قيمة الوسيط تُحسب بالعلاقة:

$$\tilde{x} := \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2}$$

وخلاف ذلك نأخذ قيمة أول ممثِّل للبيانات $x_i$ له تكرار تراكمي أكبر تماماً من $\frac{n}{2}$ كقيمة للوسيط.

**أمثلة (١, ٦, ١, ٢)**

١- لتكن لدينا البيانات الآتية:

25 31 8 42 7 33 19 5 6 14 37 29 5 19 14 8 14

والآن لتعيين الوسيط لهذه البيانات سنقوم أولاً بترتيبها تصاعدياً، فيصبح لها التوضّع الآتي للبيانات:

$x_1$ $\tilde{x} = x_9$

5 5 6 7 8 8 14 14 14 19 19 25 29 31 33 37 42

وبما أنَّ عدد هذه البيانات فردياً فإنَّ قيمة الوسيط تُحسب باستخدام العلاقة [2,13,a] فيكون لدينا:

$$\tilde{x} = x_{\frac{n+1}{2}} = x_{\frac{17+1}{2}} = x_9 = 14$$

٢- لتكن لدينا البيانات الآتية:

15 25 31 8 42 7 33 19 5 6 14 37 29 5 19 14 8 14

فعندئذ لتعيين الوسيط لهذه البيانات يجب ترتيب البيانات المُعطاة تصاعدياً، فيصبح لها التوضّع الآتي:

$x_1$ $x_9$ $\tilde{x} = 14.5$ $x_{10}$

5 5 6 7 8 8 14 14 14 15 19 19 25 29 31 33 37 42

وبما أنَّ عدد هذه البيانات زوجي فإنَّ قيمة الوسيط تُحسب باستخدام العلاقة [2-13-b] فيكون لدينا:

$$\tilde{x} = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} = \frac{x_{\frac{18}{2}} + x_{\frac{18}{2}+1}}{2} = \frac{x_9 + x_{10}}{2} = \frac{14+15}{2} = 14.5$$

٣- بالرجوع إلى المثال / ٢/ من (٢-١-٢-١) نجد أنَّ قيمة الوسيط لتلك البيانات (بعد ترتيبها تصاعدياً) يحسب باستخدام العلاقة [2-13-b]، ومن ثَّم يكون لدينا:

$$\tilde{x} = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} = \frac{x_{\frac{50}{2}} + x_{\frac{50}{2}+1}}{2} = \frac{x_{25} + x_{26}}{2} = \frac{173+174}{2} = 173.5$$

وبعد صبها في جدول توزيع تكراري نجد أنَّ الفئة الوسطية لبيانات ذلك الجدول هي الفئة الرابعة (انظر الجدول (٢, ٢))، وأمّا قيمة الوسيط لبيانات ذلك الجدول فإنَّها تحسب باستخدام العلاقة [2-14]، ومن ثَّم يكون لدينا:

$$\tilde{x} := \tilde{L} + \frac{\frac{\sum f_i}{2} - \left(\tilde{F} - \tilde{f}\right)}{\tilde{f}} \times \mathbf{C} = 170 + \frac{25-21}{15} \times 10 = 172.67$$

إنَّ الفارق بين قيمة الوسيط للبيانات الخام وقيمة الوسيط لبيانات جدول التوزيع التّكراري الممثِّل لها يعود للسبب نفسه الذي ذكرناه سابقاً بخصوص المتوسط الحسابي.

٤- لتكن لدينا مجموعتي بيانات مقدَّمة في الجدول التكراري الآتي (قيم الممثِّلين قدّمت مرتبةً تصاعدياً):

الجدول (٣،٢.*)

| رقم الممثِّل | ممثِّل البيان | التكرار $f_i$ | التكرار التراكمي للمثِّل $F_i$ | رقم الممثِّل | ممثِّل البيان | التكرار $f_i$ | التكرار التراكمي للمثِّل $F_i$ |
|---|---|---|---|---|---|---|---|
| 1 | $x_1=1$ | 3 | 3 | 1 | $z_1=5$ | 2 | 2 |
| 2 | $x_2=3$ | 5 | 8 | 2 | $z_2=7$ | 6 | 8 |
| 3 | $x_3=8$ | 3 | 11 | 3 | $z_3=11$ | 7 | 15 |
| 4 | $x_4=14$ | 4 | 15 | 4 | $z_4=13$ | 7 | 22 |
| 5 | $x_5=15$ | 4 | 19 | 5 | $z_5=17$ | 5 | 27 |
| 6 | $x_6=18$ | 6 | 25 | 6 | $z_6=25$ | 3 | 30 |
| Total | ------ | 25 | ----------------- | Total | -------- | 30 | -------------------- |

فنجد من أجل مجموعة البيانات $\mathbf{X}$ أنَّ وسيطها هو $x_4 = 14$ وذلك لأنَّ مجموع التكرارات لها عدد فردي، حيث لدينا $\frac{n+1}{2}=13$ والممثِّل $x_4$ هو أول ممثِّل بيانات له تكرار تراكمي أكبر أو يساوي 13. أمّا من أجل مجموعة البيانات $\mathbf{Z}$ فنجد أنَّ مجموع التكرارات لها عدد زوجي، وللممثِّل $z_3$ تكرار تراكمي يساوي $\frac{n}{2}$، ومن ثمَّ قيمة الوسيط لمجموعة البيانات $\mathbf{Z}$ هو:

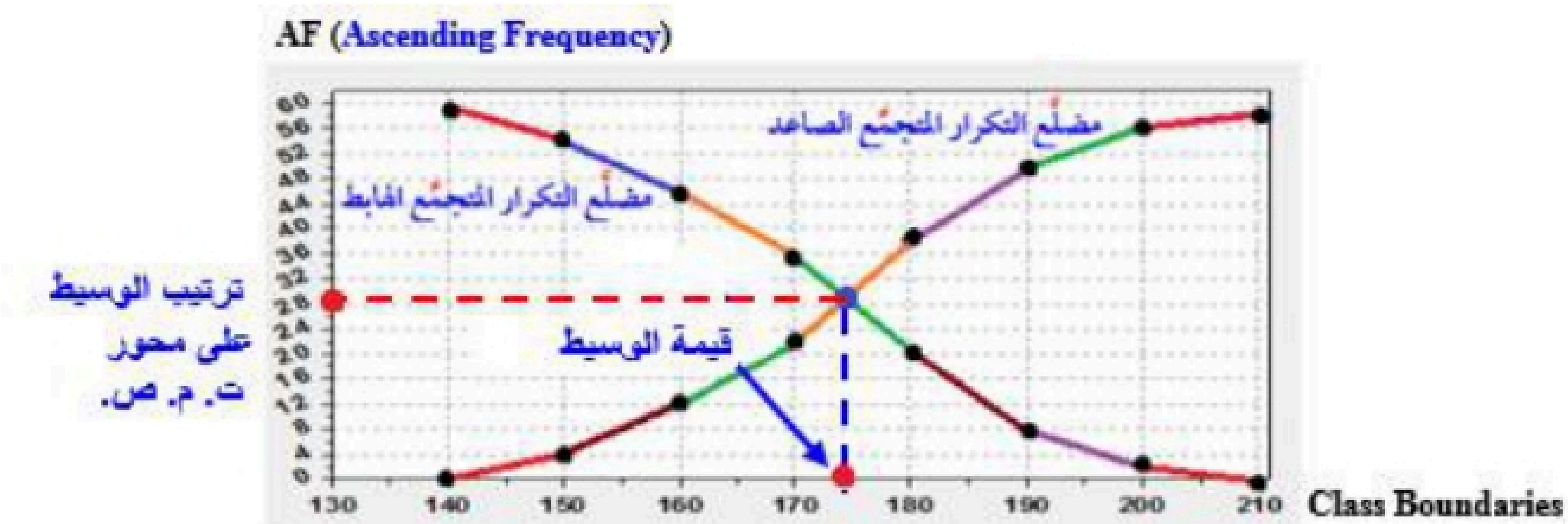
$$\tilde{x} = \frac{x_3+x_4}{2} = \frac{11+13}{2} = 12$$

### (٢،٦،١،٢) ملاحظات

١- إذا فقدت إحدى أو بعض البيانات المعلوم ترتيبها بين البيانات وغير الواقعة في وسط البيانات المرتّبة فإنَّ الوسيط قابل للتطبيق والحساب.

٢- يُعبَّر عن الوسيط هندسياً على أنَّه القيمة $x$ على محور الفئات التي إذا رسم عندها عمود رأسي فإنَّه سيقسم المدرَّج التّكراري (أو المضلَّع التّكراري) إلى قسمين متساويين (وليس متماثلين بالضرورة).

٣- يمكن تعين المنوال هندسياً من خلال إسقاط عمود على محور الفئات من نقطة تقاطع مضلَّعي التكرار المتجمِّع الصاعد والهابط، فلو عدنا إلى الفقرة (٣،٥،١) فإنَّ قيمة الوسيط لبيانات جدول التوزيع التكراري (١٥،١) موضَّحةً في الشكل الآتي.




الشكل (٣,٢)

**(٣,٦,١,٢) أمثلة**

١ – لتكن لدينا البيانات الآتية التي تعرض لنا درجات الحرارة في مدينة X خلال أحد أيام شهر شباط (فبراير) مسجلةً كل ساعة بدءًا من الساعة 01:00 صباحاً:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 | 7 | 8 | 6 | 6 | 7 | 8 | 9 | 10 | 13 | 18 | 7 |
| 18 | 18 | 17 | 7 | 17 | 16 | 14 | 13 | 11 | 10 | 9 | 7 |

فنجد أنَّ عدد البيانات زوجياً، وأنَّ الترتيب التصاعدي لهذه البيانات له العرض الآتي:

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 6 | 6 | 7 | 7 | 7 | 7 | 7 | 8 | 8 | 8 | 9 | 9 | 10 | 10 | 11 | 13 | 13 | 14 | 16 | 17 | 17 | 18 | 18 | 18 |
| | | | | | | | | | | | $x_{11}$ | $x_{12}$ | | | | | | | | | | | |

ومن ثمَّ تكون قيمة الوسيط لهذه البيانات هي:

$$\tilde{x} = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} = \frac{x_{12} + x_{13}}{2} = \frac{9+10}{2} = 9.5$$

فلو كانت لدينا بعض البيانات مفقودة (على سبيل المثال، والتي سنشير إليها بـ ?) المعلوم ترتيبها وغير الواقعة في الوسط، فإنَّنا سنلاحظ أنَّ قيمة الوسيط تبقى قابلة للحساب.

| | | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 6 | 6 | 7 | 7 | ? | 7 | 7 | 8 | ? | 8 | 9 | 9 | 10 | 10 | 11 | 13 | ? | 14 | 16 | 17 | 17 | 18 | 18 | ? |

٢ – بالعودة إلى مثال جمع التبرعات من أحد الفصول الدراسيّة بمدرسة ما لصالح إحدى الجمعيات الخيرية، وقد لاحظنا سابقاً تأثّر قيمة المتوسط الحسابي للتبرعات بالقيمة المتطرفة 100، ولكن لو قمنا بترتيب هذه البيانات تصاعدياً:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 3 | 3 | 4 | 4 | 4 | 4 | 4 | 4 |
| 4 | 4 | 5 | 5 | 5 | 5 | 5 | 5 | 5 | 5 |
| 6 | 6 | 6 | 7 | 7 | 7 | 7 | 7 | 100 | 7 |

ومن ثمَّ نجد أنَّ الوسيط لهذه التبرعات هو:

$$\tilde{x} = \frac{x_{15} + x_{16}}{2} = \frac{5+5}{2} = 5$$

فنجد أنَّ 50% من المتبرعين لم يتبرعوا بأكثر من 5 ريالات، وهذا يعني أنَّ هذا المقياس للنَّزعة المركزيَّة المتمثِّل بالقيمة 5 لم يتأثر بالقيمة المتطرفة 100.

٣- لتكن لدينا البيانات الآتية التي تعرض لنا درجات مجموعة من الطلاب (عددهم 47 طالباً) في اختبار منتصف الفصل لمقرّر الإحصاء (الدرجة القصوى 20):

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 13.5 | 18 | 8.5 | 13.5 | 11.5 | 8 | 13 | 9 | 17.5 | 17 | 14.5 | 8 |
| 13.5 | 7 | 11 | 7.5 | 6 | 14 | 15.5 | 14 | 13.5 | 14.5 | 16 | 11.5 |
| 8 | 17.5 | 18 | 20 | 15.5 | 9.5 | 1 | 5.5 | 13.5 | 9.5 | 12.5 | 7 |
| 9.5 | 15 | 5 | 8 | 11.5 | 18.5 | 7 | 10 | 8 | 11.5 | 7 | |

وبترتيب هذه البيانات تصاعدياً يصبح لدينا العرض الآتي للبيانات السّابقة:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 5 | 5.5 | 6 | 7 | 7 | 7 | 7.5 | 8 | 8 | 8 | 8 |
| 8.5 | 9 | 9.5 | 9.5 | 9.5 | 10 | 11 | 11.5 | 11.5 | 11.5 | 12.5 | 13 |
| 13 | 13.5 | 13.5 | 13.5 | 13.5 | 14 | 14 | 14.5 | 14.5 | 14.5 | 15 | 15.5 |
| 15.5 | 16 | 17 | 17.5 | 17.5 | 17.5 | 17.5 | 18 | 18 | 18.5 | 20 | |

وبحساب قيمة الوسيط لهذه البيانات نجد أنَّ:

$$\tilde{x} = x_{\frac{n+1}{2}} = x_{\frac{47+1}{2}} = x_{24} = 13$$

وهكذا نجد أنَّ 50 % من الطلاب لهم درجات اختبار أعلى من 13 درجة.

لقد ذكرنا سابقًا أنَّه من الممكن لمجموعة بيانات أن تنزع إلى أكثر من قيمة، بمعنى أن مجموعة البيانات تتجزأ بحيث كل مجموعة جزئية منها تنزع إلى قيمة مختلفة عن الأخرى، وقد يحصل هذا عادة عندما تكون البيانات متأتية من مصادر مختلفة. إنَّ التعريف الآتي يقدّم لنا تعريف هذا النوع من مقاييس النّزعة المركزيَّة، حيث سنلاحظ أنَّ هذا المقياس يتميَّز بسلوك مختلف عن المقاييس السّابقة.

**(٢, ١, ٧) تعريف (المنوال Mode)**

سنرمز للمنوال بـ $\hat{x}$ من أجل العينات وبـ $\hat{\mu}$ من أجل المجتمعات، ولتعريفه سنناقش الحالتين الآتيتين:

**١- من أجل البيانات الخام:**

بفرض أنَّ $x_1$، $x_2$، ...، و $x_n$ بيانات خام مُعطاة (لعيِّنة أو لمجتمع الفارق هو استخدام الدليل $n$ أو $N$ فقط)، عندئذ يُعرَّف المنوال على أنَّه القيمة الأكثر تكراراً بين البيانات.

**٢- من أجل البيانات المجمَّعة:**

بفرض أنَّ لدينا بيانات مُعطاة من خلال جدول التوزيع التكراري (١ , ٢)، فعندئذ تُحسب قيمة المنوال بوساطة العلاقة الآتية:

$$\hat{x} = \hat{L} + \frac{D_1}{D_1 + D_2}\mathbf{C} \qquad [2,15]$$

علماً أنَّ: $\hat{L}$ هو الحد الأدنى للفئة المنواليّة، $D_1$ هو الفرق بين تكرار الفئة المنواليّة وتكرار الفئة السّابقة لها مباشرةً، $D_2$ هو الفرق بين تكرار الفئة المنواليّة وتكرار الفئة اللاحقة بها مباشرةً، واخيراً $\mathbf{C}$ هو سعة الفئة المنواليّة (هنا لم نرفق دليلًا لـ $\mathbf{C}$ لأنَّه لدينا جميع الفئات متساوية السعة). علماً أنَّ الفئة المنواليّة هي تلك الفئة التي تكرارها أكبر من تكرار الفئة السّابقة لها واللاحقة بها مباشرةً، وعلى ألاّ تكون هذه الفئة طرفية. أي إن الفئة الأولى والأخيرة في التوزيع التّكراري لا يُنظر إليهما كفئات منواليّة.

أمّا إذا كانت البيانات مقدَّمة من خلال جدول تكراري كما في (١ , ٢) (مع مراعاة الملاحظة المرفقة به)، فعندئذ كل قيمة ممثِّل للبيانات يقابلها أكبر قيمة تكرار ستكون منوالاً للبيانات (بالطبع إذا كانت جميع قيم التكرارات متساوية فعندئذ ليس للبيانات منوال).

**(١, ٧, ١, ٢) أمثلة**

١- لتكن لدينا البيانات الآتية:

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | 4 | 15 | 10 | 7 | 9 | 7 | 11 | 7 | 6 | 2 | 11 | 3 | 11 | 8 | 11 | 5 |

فنجد أنَّ القيمة الأكثر تكراراً هي 11، ومن ثَّم المنوال لهذه البيانات يساوي 11.

٢- بالرجوع إلى المثال / ٢/ من (١, ٢, ١, ٢) نجد أنَّ القيمة الأكثر تكراراً هي 175 (تكرارها يساوي 5)، ومن ثَّم المنوال لتلك البيانات يساوي 175. أمّا بعد صبّ البيانات المُعطاة في جدول توزيع تكراري فنجد أنَّ الفئة المنواليّة لبيانات جدول التوزيع التّكراري (٢, ٢) هي الفئة الثالثة (انظر الجدول (٢, ٢))، ومن ثَّم باستخدام العلاقة [2,15] تكون قيمة المنوال لبيانات ذلك الجدول تساوي:

$$\hat{x} = \hat{L} + \frac{D_1}{D_1 + D_2}\,\mathbf{C} = 160 + \frac{14}{14+2}\times 10 = 168.75$$

إنَّ الفارق بين قيمة المنوال للبيانات الخام وبيانات جدول التوزيع التّكراري الممثِّل لها يعود للسبب نفسه الذي ذكرناه سابقاً بخصوص المتوسط الحسابي أيضاً.

٣- بالرجوع إلى المثال / ٣/ من (١, ٦, ١, ٢) نجد أنَّ مجموعة البيانات **X** تملك منوالاً وحيداً هو $x_5 = 15$، وأمّا مجموعة البيانات **Z** فنجد أنّها تملك منوالين هما $z_3 = 11$ و $z_4 = 13$ حيث لكلٍ منها التكرار الأعظمي 7.

**(٢, ٧, ١, ٢) ملاحظات**

١- إنَّ المنوال قد لا يكون موجوداً، وهذا يوافق الحالة التي تكون فيها جميع البيانات الخام مختلفة بعضها عن لبعض، من قبيل:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 8 | 5 | 9 | 4 | 7 | 6 | 2 | 11 | 3 | 15 | 10 |

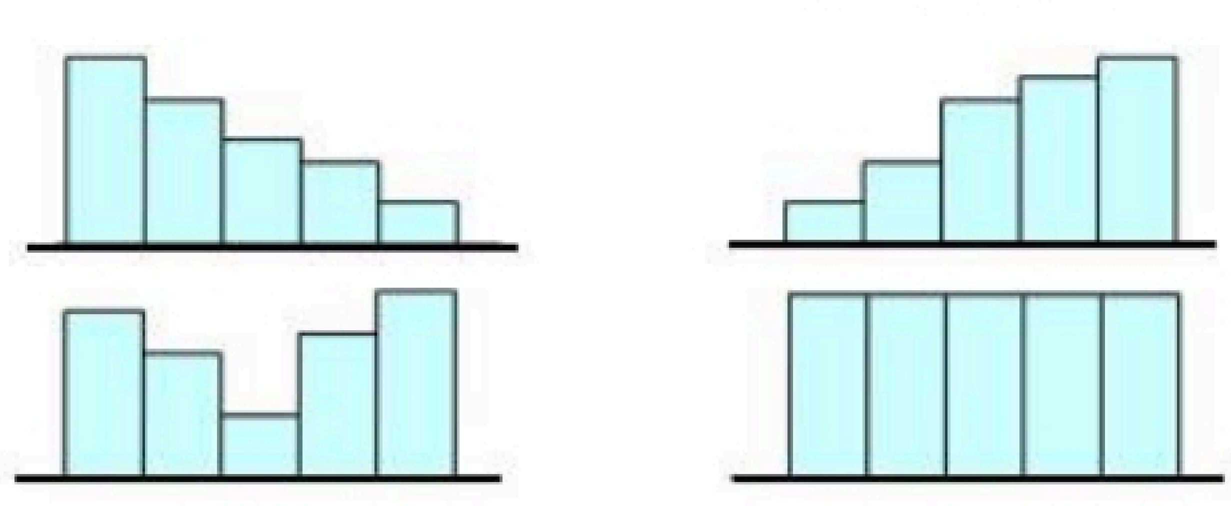
أو أنَّ تكون جميع البيانات متساوية، من قبيل:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 | 7 |

أو أنَّ يكون لجميع البيانات التكرار نفسه، من قبيل:

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 8 | 5 | 8 | 4 | 5 | 6 | 4 | 6 | 0 | 0 | 8 | 5 | 4 | 6 |

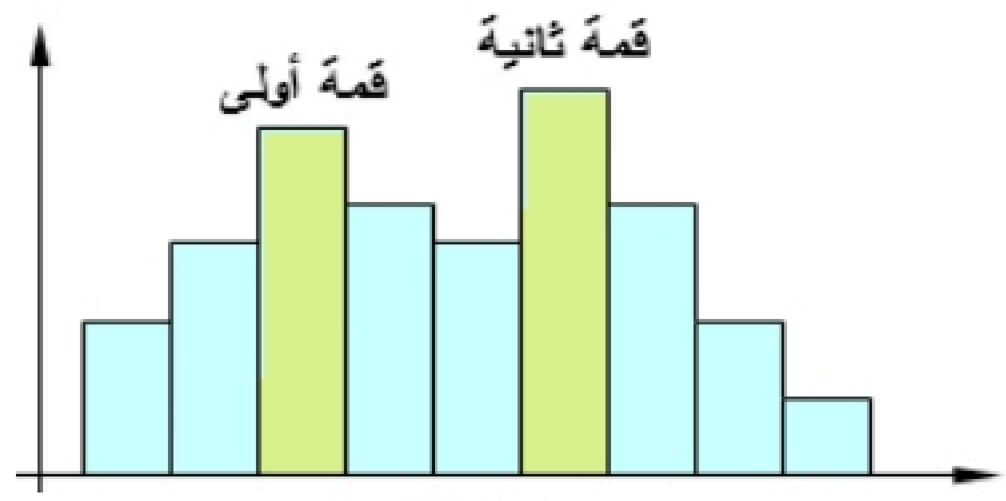
وأمّا من أجل التوزيعات التكرارية فإنَّ عدم وجود المنوال يوافق الحالة التي يكون فيها شكل المدرَّج التّكراري (وقيا ساً عليه تناقش بقية التمثيلات) للبيانات المجمَّعة أحد العروض البيانية الآتية:



٢- حتى في حال وجود المنوال، فإنَّه قد لا يكون وحيداً، وهذا يوافق الحالة التي يتساوى فيها التكرار الأعظمي لقيمتين من البيانات على الأقل، فيكون لدينا في هذه الحالة منوالين على الأقل، فعلى سبيل المثال، لو أخذنا البيانات الآتية:

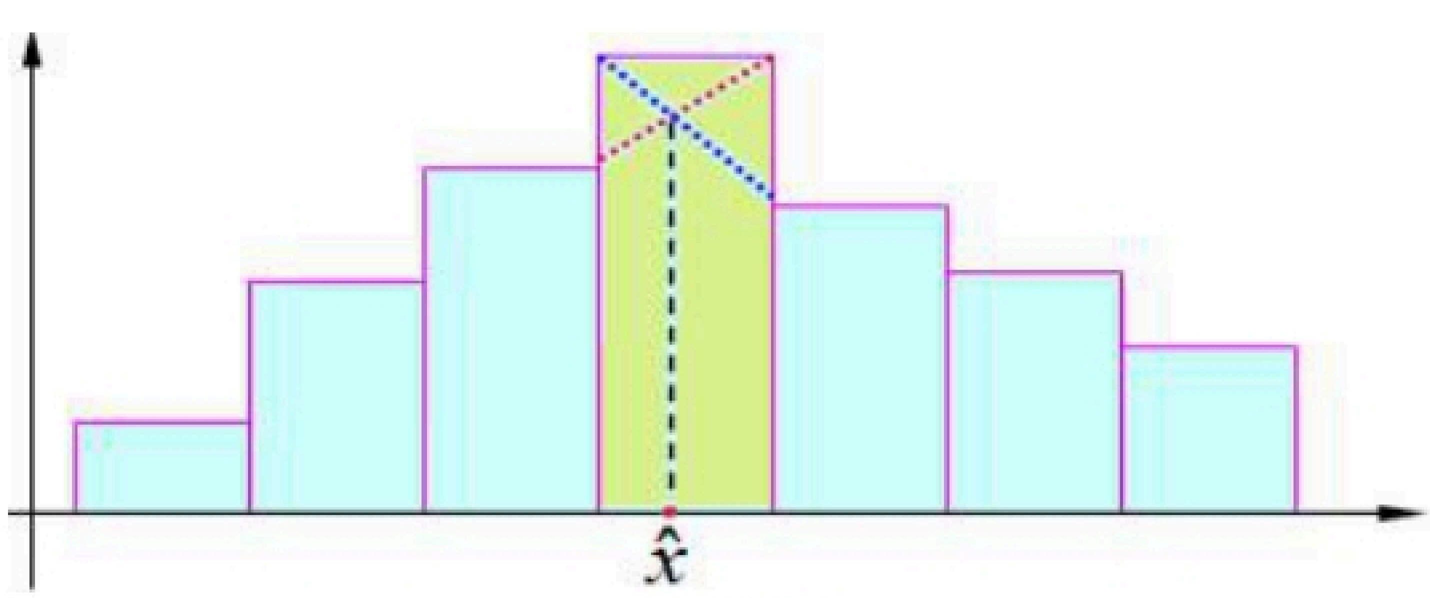
| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 | 5 | 7 | 6 | 8 | 11 | 3 | 10 | 10 | 7 | 9 | 7 | 11 | 11 | 8 |

فنلاحظ وجود ثلاثة مناويل لهذه البيانات هي 7، 8 و11 (حيث لكل منها ثلاثة تكرارات)، و أمّا من أجل البيانات المجمّعة فيكون لشكل المدرّج التّكراري (وقياساً عليه تناقش بقية التمثيلات) للبيانات المجمَّعة قمتين داخليتين على الأقل، والشّكل الآتي يوضِّح ذلك (مدرّج تكراري ذو منوالين).




الشكل (٢, ٤)

٣- إذا كان المدرَّج التّكراري مُعطى فإنَّه يمكن تعين المنوال هندسياً، وذلك من خلال إسقاط عمود على محور الفئات من نقطة تقاطع المستقيمين الواصلين بين النهاية اليمنى لقمة الفئة المنواليّة مع النهاية اليمنى لقمَّة الفئة السّابقة لها، وكذلك بين النهاية اليسرى لقمَّة الفئة المنواليّة مع النهاية اليسرى لقمَّة الفئـة اللاحقة بها كما يوضّحه الشّكل الآتي.




الشكل (٢, ٥)

وهذه الطريقة لتعيين المنوال تحتاج إلى دقة في الرسم واستخدام مقاييس دقيقة، ولذلك لا تستخدم إلاَّ عند الضرورة.

### (٢, ١, ٨) مقارنة بين صفات المتوسط، والوسيط، والمنوال

١- المتوسط هو أكثر مقاييس النَّزعة المركزيَّة استخداماً في مجال الإحصاء، ويعود ذلك لسهولة حسابه وتعريفه بعلاقة رياضية بسيطة، وكذلك لكونه يأخذ بالحسبان جميع القياسات التي تخضع للدراسة والبحث، كما أنَّه يخضع للعمليّات الجبرية المعروفة، ومن ميزاته أنَّه إذا عُلمت قيمته فإنَّه يمكن حساب مجموع البيانات إذا كان عددها معلوماً، أو حساب عدد البيانات إذا كان المجموع لها معلوماً، وهذه الخاصية غير متوفرة في الوسيط والمنوال، وإضافة لما سبق فإنَّنا نلاحظ أنَّه إذا أضفنا إلى البيانات أي عدد من القيم المساوية للمتوسط فإنَّ قيمة هذا المتوسط لا تتغيّر وذلك لأنَّ المتوسط يُمثِّل مركز ثقل البيانات أو نقطة اتزان المدرَّج التّكراري، وأخيراً يُفضّل استخدامه في الاستنتاج الإحصائي ذلك أنَّ توزيعاته معلومة في الإحصاء (كما سنرى ذلك لاحقاً في الفصل التاسع).

**أمّا أهمّ عيوب المتوسط الحسابي فهي:**

أ- تأثره بالقيم المتطرفة، وللتخلص من هذا العيب يمكن استخدام ما يُعرف باسم المتوسط الحسابي المشذَّب Trimmed mean (TrMean) الذي يقوم على ترتيب البيانات أولاً، ومن ثمَّ حذف أول 5% وآخر 5% من البيانات (مقرّبةً إلى أقرب عدد صحيح)، وبعد ذلك حساب المتوسط لقيم البيانات المتبقية.

ب- يصبح عديم التطبيق في حال سقوط أو فقدان إحدى قيم البيانات أو بعضها.

٢- الوسيط سهل التعريف والحساب، ولا يتأثر بالقيم المتطرّفة، ولا يعتمد على جميع القيم دائمًا، فتغيّر قيمة من القيم غير الواقعة في الوسط لا يؤثّر في قيمة الوسيط، كذلك يستعمل الوسيط في البيانات التي يُعرف ترتيبها ولا تعرف قيمها، وكذلك في البيانات الناقصة كأن تكون إحدى قيم البيانات غير الواقعة في الوسط قد سقطت لسبب ما ولكن معلوم ترتيبها.

**أمّا أهمّ عيوب الوسيط فهي:**

أ - انخفاض مستوى دقته مقارنة بالمتوسط لأنَّه يعتمد على قيمة واحدة أو اثنتين على الأكثر من قيم البيانات فقط.

ب- عدم إمكانية تطبيقه في حال فقدان بعض البيانات المرتبة الواقعة في الوسط.

٣- المنوال هو أقل هذه المقاييس الثلاثة استخداماً، وفي البيانات القليلة العد يعدُّ غير ذي جدوى عمليًّا (هذا إن وجد أصلاً)، وأمّا في حالة البيانات كبيرة العدد، والجداول التكرارية فيمكن الاستفادة منه لكون تأثره بتغيُّر بعض قيم البيانات قد لا يكون كبيراً بالقدر الذي يلغي استخدامه، كذلك من فوائده أنَّه يمكن استخدامه في حالة البيانات النوعيّة مثل: معرفة ألوان عيون فئة من الناس، فيُقال إنَّ اللون المنوالي هو اللون البنّي، ولكن من عيوبه الكبيرة أنَّه قد لا يكون موجوداً، وفي حال وجوده قد لا يكون وحيداً، وهذا يعني أنَّه لا يمكن تعريفه بشكل وحيد كمقياس للنَّزعة المركزيَّة.

**(١ ,٨ ,١ ,٢) ملاحظات**

١- بالرغم من أنَّ المتوسط يُنظر إليه في كثير من الحالات كأفضل مقياس من مقاييس النَّزعة المركزيَّة، إلاَّ أنَّه يُفضّل استخدامه عندما يكون شكل توزيع البيانات متماثلاً على وجه التقريب، وعندما يكون الاهتمام منصباً على القيمة العددية التي تأخذ جميع القياسات بالحسبان وليس الحصول على قيمة نموذجية مُمثِّلة لها فقط، وكذلك عندما تكون جميع البيانات معلومة (أي لا يوجد فيها بيانات مفقودة).

٢- يُفضّل استخدام الوسيط إذا كان اهتمامنا منصباً على إيجاد قيمة مُمثِّلة لمركز ثقل البيانات بدلاً من الاهتمام بالمجموع الكلي، وعندما يكون التوزيع ملتوياً، وكذلك عندما تكون لدينا بعض البيانات المفقودة المعلوم ترتيبها وغير الواقعة في وسط البيانات بعد ترتيبها (تصاعديًا أو تنازليًا).

٣- إنَّ استخدام المنوال قليل جداً ونادراً ما يُلجأ إليه، ولكن في حال فقدان بعض البيانات الواقعة في وسط البيانات بعد ترتيبها فإنَّ للمنوال دور جيد في تعيين مقياس النَّزعة المركزيَّة للبيانات.

٤- وجد تجريبياً أنَّ المنوال يساوي ثلاثة أمثال الوسيط مطروحاً منه ضعف المتوسط، أي إنَّ:

$$\hat{x} = 3\tilde{x} - 2\bar{x}$$

[2,16]

## (٢,٢) الربيعيات

## Quartiles

لقد ناقشنا سابقاً بعضاً من مقاييس الموضع المهمّة التي تمثّلت بمقاييس النَّزعة المركزيَّة، ولكن هناك مقاييس موضع أخرى لا تنتمي لمقاييس النَّزعة المركزيَّة. سنذكر منها نوعين هما الرُّبَيْعيّات والمئينات فقط.

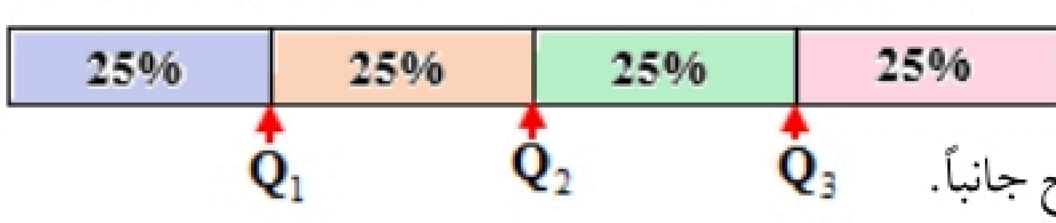
### (١ , ٢ , ٢) تعريف ( الرُّبَيْعيّات)

لتكن لدينا $x_1$، $x_2$، ...و $x_n$ مجموعة من البيانات المرتّبة تصاعدياً، فعندئذ:

١ – القيمة التي يقع قبلها 25% وبعدها 75% من البيانات المرتّبة تُدعى **الرُّبَيْعيّ الأول** First Quartile (**أو الرُّبَيْعيّ الأدنى Lower Quartile**)، ويُرمز له بـ $\mathbf{Q}_1$.

٢ – القيمة التي يقع قبلها 50% وبعدها 50% من البيانات المرتّبة تُدعى **الرُّبَيْعيّ الثاني** Second Quartile، ويُرمز له بـ $\mathbf{Q}_2$، وهنا يُلاحظ أنَّ الرُّبَيْعيّ الثاني هو الوسيط نفسه، أي أنَّ $\mathbf{Q}_2 = \tilde{x}$.

٣– القيمة التي يقع قبلها 75% وبعدها 25% من البيانات المرتّبة تُدعى **الرُّبَيْعيّ الثالث** Third Quartile (**أو الرُّبَيْعيّ الأعلى Upper Quartile**)، ويُرمز له بـ $\mathbf{Q}_3$.




هكذا يمكننا تقديم العرض الهيكلي لهذه الرُّبَيْعيّات كما موضّح جانباً.

### (٢ , ٢ , ٢) تعيين الرُّبَيْعيّات

في الواقع توجد طرائق عديدة لتعيين الرُّبَيْعيّات، وتتفاوت نتائجها وفقاً لهذه الطريقة أو تلك، وهذا بدوره قد يجعل نتائج حساب الرُّبَيْعيّات تختلف من برنامج إحصائي إلى آخر. في كتابنا هذا سنقدّم إحدى هذه الطرائق التي تُعدّ من أكثر الطرائق دقةً، وهي مستخدمة في بعض البرامج الإحصائية.

لتعيين الرُّبَيْعيّات سنميِّز بين حالتين:

**١– من أجل البيانات الخام:**

لتكن $x_1$، $x_2$، ...و $x_n$ بيانات خام مرتبة تصاعدياً (**لعيِّنة أو لمجتمع الفارق هو استخدام الدليل $n$ أو $N$ فقط**)، ولنضع بالتعريف:

$$q_r := \frac{r\ (n+1)}{4} \quad ; r = 1, 2, 3 \qquad [2,17]$$

والذي يُدعى **رتبة Rank** الرُّبَيْعيّ $r$ (**وهذه القيمة تحدِّد موضع الرُّبَيْعيّ بين البيانات المرتّبة**)، فعندئذ تُعيَّن قيمة الرُّبَيْعيّ $\mathbf{Q}_r$ مع $r = 1, 2, 3$ على النحو الآتي:

بفرض أنَّ $k$ هو الجزء الصحيح من العدد $q_r$، وأنَّ الباقي من هذا العدد يساوي $s$، فحينئذ يُحسب الرُّبَيْعيّ $\mathbf{Q}_r$ من خلال العلاقة:

$$\mathbf{Q}_r = x_k + s\ (x_{k+1} - x_k) \quad ; r = 1, 2, 3 \qquad [2,18]$$

لاحظ هنا أنَّه إذا كان الباقي $s = 0$ أو كان $x_{k+1} = x_k$، فعندئذ سيكون لدينا $\mathbf{Q}_r = x_k$.

**٢– في حالة البيانات المجمَّعة:**

بفرض أنَّه لدينا بيانات مُجَمَّعة في جدول توزيع تكراري كما في الجدول (١ , ٢)، ولنضع بالتعريف:

$$q_r := \frac{r \cdot \sum f_i}{4} \qquad [2,19]$$

فـ عندئذ يُعيَّن الرُّبَيْعيّ ذو الرقم $r$ لبيانات هذا الجدول من خلال العلاقة الآتية:

$$\mathbf{Q}_r := L_{\mathbf{Q}_r} + \frac{q_r - \left(F_{\mathbf{Q}_r} - f_{\mathbf{Q}_r}\right)}{f_{\mathbf{Q}_r}} \times \mathbf{C} \qquad ; \ r = 1,2,3 \qquad [2,20]$$

علماً أنَّ $L_{\mathbf{Q}_r}$ هو الحد الأدنى لفئة الربيعي $\mathbf{Q}_r$، وأنَّ فئة هذا الرُبَيْعيّ هي أول فئة تكرارها المتجمِّع الصاعد أكبر أو يساوي $q_r$، $f_{\mathbf{Q}_r}$ هو تكرار فئة الرُبَيْعيّ $r$، و $F_{\mathbf{Q}_r}$ هو التكرار المتجمّع الصاعد لفئة الرُبَيْعيّ $r$، وأخيراً $\mathbf{C}$ هي سعة الفئة للربيعي $r$ (هنا لم نرفق دليلًا لـ $\mathbf{C}$ لأنّنا نتعامل مع جداول توزيع تكرارية فئاتها متساوية السعة).

في حال كانت البيانات مقدَّمة من خلال جدول تكراري، فعندئذ نعيد تنظيم الجدول بحيث تصبح قيم الممثِّلين في الجدول التكراري (٢،١) مرتّبةً تصاعدياً. بعد ذلك نفترض أنَّ البيانات الأصل هي $y_1$ ، $y_2$، ... و $y_n$ مع $n = \sum f_i$ و $y_1 < y_2 < ... < y_n$، ومن ثَمَّ تُحسب رتبة الرُبَيْعيّ $r$ بوساطة العلاقة [2,17] لهذه البيانات، فإن كانت قيمة $q_r$ عدداً صحيحاً ووجدت قيمة ممثِّل $x_i$ في الجدول التكراري لها تكرار تراكمي يساوي $q_r$ فإنَّنا نأخذ تلك القيمة كقيمة لـ المُرَبْيعيّ $r$، وخلاف ذلك نأخذ أول قيمة لممثِّل $x_i$ في الجدول التكراري لها تكرار تراكمي أكبر من $q_r$ كقيمة لـ المُرَبْيعيّ $r$ . أمّا إذا كانت قيمة $q_r$ عدداً غير صحيح، فعندئذ نطبِّق على البيانات المرتَّبة $y_1$ ، $y_2$، ... و $y_n$ العلاقة الآتية:

$$\mathbf{Q}_r = y_\ell + s\,(y_{\ell+1} - y_\ell) \ \ ; r = 1, 2, 3$$

علماً أنَّ $\ell$ هنا هو الجزء الصحيح من $q_r$ و $y_\ell$ هي أول قيمة ممثِّل $x_i$ في الجدول التكراري لها تكرار تراكمي أكبر أو يساوي $\ell$، و $y_{\ell+1}$ هي القيمة $x_i$ نفسها إذا كان $F_i > \ell$، وخلاف ذلك (أي أنَّ $F_i = \ell$) تكون $y_{\ell+1}$ هي $x_{i+1}$ من الجدول التكراري.

### (٢، ٢، ٢، ١) أمثلة

١- تمَّ معايرة المادة الفعّالة لعيِّنة عشوائية بحجم $n = 20$ كبسولة سُحِبت من إنتاج مصنع لكبسولات مضاد حيوي عيار 500 mg، فكانت لدينا البيانات الآتية بعد ترتيبها تصاعدياً:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **480** | **482** | **483** | **485** | **490** | **492** | **495** | **495** | **500** | **500** |
| **505** | **510** | **510** | **515** | **525** | **525** | **535** | **537** | **549** | **550** |

ولنقم بتعيين الرُبَيْعيّات الثلاثة $\mathbf{Q}_1$، $\mathbf{Q}_2$ و $\mathbf{Q}_3$ ونقدِّم تفسيراً لها.

**العمل:** إنَّ رتبة الرُبَيْعيّ الأول، والثاني، والثالث على الترتيب هي:

$$q_1 = \frac{n+1}{4} = \frac{20+1}{4} = 5.25 \ \ \& \ \ q_2 = \frac{2(n+1)}{4} = \frac{20+1}{2} = 10.5 \ \ \& \ \ q_3 = \frac{3(n+1)}{4} = \frac{3(20+1)}{4} = 15.75$$

ومن ثَمَّ تكون قيم الرُبَيْعيّات الثلاثة $\mathbf{Q}_1$، $\mathbf{Q}_2$ و $\mathbf{Q}_3$ على الترتيب هي:

$$\mathbf{Q}_1 = x_5 + 0.25(x_6 - x_5) = 490 + 0.25\ (492 - 490) = 490.5$$

$$\mathbf{Q}_2 = x_{10} + 0.5\ (x_{11} - x_{10}) = 500 + 0.5\ (505 - 500) = 502.5$$

$$\mathbf{Q}_3 = x_{15} + 0.75\ (x_{15} - x_{16}) = 525 + 0.75\ (525 - 525) = 525$$

وهذا يعني أنَّ وفقاً لمعطيات العيِّنة لدينا 25% من المنتج يحوي على كمّيّة **أقل** من 490.5 mg، و**نصف** المنتج يحوي على كمّيّة أقل من 502.5 mg، واخيراً 25% من المنتج يحوي على كمّيّة **أكثر** من 525 mg.

٢- لتكن لدينا بيانات مجمَّعة في جدول التوزيع التّكراري الآتي، والتي تمثِّل نتائج الاختبار النهائي لـ 48 طالباً في مقرَّر الإحصاء:

الجدول (٤ , ٢)

| رقم الفئة | الحدود الفعليّة للفئة | مركز الفئة $x_i$ | تكرار الفئة $f_i$ | التكرار المتجمّع الصاعد للفئة $F_i$ | رقم الفئة |
|---|---|---|---|---|---|
| **1** | **25 → 30** | **27.5** | **5** | **5** | **1** |
| **2** | **30 → 35** | **32.5** | **9** | **14** | **2** |
| **3** | **35 → 40** | **37.5** | **16** | **30** | **3** |
| **4** | **40 → 45** | **42.5** | **8** | **38** | **4** |
| **5** | **45 → 50** | **47.5** | **5** | **43** | **5** |
| **6** | **50 → 55** | **52.5** | **3** | **46** | **6** |
| **Total** | ------------- | ---- | **48** | ------------------- | **Total** |

ولنقم بتعيين الرُّبَيْعيّات الثلاثة $\mathbf{Q}_1$، $\mathbf{Q}_2$ و $\mathbf{Q}_3$ ونقدِّم تفسيراً لها.

**العمل**: إنَّ رتبة الرُّبَيْعيّ الأول، والثاني، والثالث على الترتيب هي:

$$q_1 = \frac{48}{4} = 12 \quad \& \quad q_2 = \frac{2 \times 48}{4} = 24 \quad \& \quad q_3 = \frac{3 \times 48}{4} = 36$$

ومن ثَمَّ تكون قيمة الرُّبَيْعيّ الأول، والثاني، والثالث على الترتيب هي:

$$\mathbf{Q}_1 = L_{\mathbf{Q}_1} + \frac{q_1 - \left(F_{\mathbf{Q}_1} - f_{\mathbf{Q}_1}\right)}{f_{\mathbf{Q}_1}} \times \mathbf{C} = 30 + \frac{12 - (14 - 9)}{9} \times 5 = 33.89$$

$$\mathbf{Q}_2 = L_{\mathbf{Q}_2} + \frac{q_2 - \left(F_{\mathbf{Q}_2} - f_{\mathbf{Q}_2}\right)}{f_{\mathbf{Q}_2}} \times \mathbf{C} = 35 + \frac{24 - (30 - 16)}{16} \times 5 = 38.125$$

$$\mathbf{Q}_3 = L_{\mathbf{Q}_3} + \frac{q_3 - \left(F_{\mathbf{Q}_3} - f_{\mathbf{Q}_3}\right)}{f_{\mathbf{Q}_3}} \times \mathbf{C} = 40 + \frac{36 - (38 - 8)}{8} \times 5 = 43.75$$

ومن ثَمَّ وفقاً لمعطيات هذه العينة فإنَّ 75% من الطلاب الذين تقدّموا إلى الاختبار حصلوا على درجات أعلى من 33.89 درجة، وأنَّ نصفهم حصلوا على درجات فوق 39 درجة (ومثلهم دون 39 درجة)، وأخيراً فإنَّ 75% منهم حصلوا على درجات أقل من 44.125 درجة.

**٣**- بالرجوع إلى المثال / ٣/ من (١ , ٦ , ١ , ٢) . لنقم بتعيين الرُّبَيْعيّات الثلاثة $\mathbf{Q}_1$، $\mathbf{Q}_2$ و $\mathbf{Q}_3$ لمجموعة البيانات $\mathbf{X}$ .

ولنقم بتعيين الرُّبَيْعيّات الثلاثة $\mathbf{Q}_1$، $\mathbf{Q}_2$ و $\mathbf{Q}_3$، فنجد أنَّ:

$$q_1 = \frac{26}{4} = 6.5 \quad \& \quad q_2 = \frac{54}{4} = 13.5 \quad \& \quad q_3 = \frac{78}{4} = 19.5$$

ومن ثَمَّ نجد أنَّ جميع قيم $q_r$ هي أعداد غير صحيحة، وبالتالي يكون لدينا:

- من أجل $\mathbf{Q}_1$ لدينا الجزء الصحيح من $q_r$ هو $\ell = 6$ ومن ثَمَّ تكون قيمة $y_\ell = x_2 = 3$، وبما أنَّ $F_2 = 8 > \ell = 6$ فإنَّ $y_{\ell+1} = x_2 = 3$ أيضاً، وبالتالي يكون لدينا:

$$\mathbf{Q}_1 = y_6 + 0.5(y_7 - y_6) = x_2 + 0.5(x_2 - x_2) = x_2 = 3$$

- من أجل $\mathbf{Q}_2$ لدينا الجزء الصحيح من $q_r$ هو $\ell = 13$ ومن ثَّم تكون قيمة $y_\ell = x_4 = 14$ ، وبما أنَّ $F_4 = 15 > \ell = 13$ فإنَّ $y_{\ell+1} = x_4 = 14$ أيضاً، وبالتالي يكون لدينا:

$$\mathbf{Q}_1 = y_{13} + 0.5(y_{14} - y_{13}) = x_4 + 0.5(x_4 - x_4) = x_4 = 14$$

- من أجل $\mathbf{Q}_3$ لدينا الجزء الصحيح من $q_r$ هو $\ell = 19$ ، ومن ثَّم تكون قيمة $y_\ell = x_5 = 15$ ، وبما أنَّ $F_5 = \ell = 19$ فإنَّ $y_{\ell+1} = x_6 = 18$ ، وبالتالي يكون لدينا:

$$\mathbf{Q}_3 = y_{19} + 0.5(y_{20} - y_{19}) = x_5 + 0.5(x_6 - x_5) = 15 + 0.5(18 - 15) = 16.5$$

**(٢,٢,٢,٢) ملاحظات**

١- تجدر الإشارة هنا إلى أنَّه من الممكن أن تكون قيمة الرُّبَيْعيّ موجودةً بين القيم المرتّبة المُعطاة.

٢- لقد لاحظنا أنَّ القيمة $\mathbf{Q}_2$ هي الوسيط، ومن ثَّم $\mathbf{Q}_2$ هو مقياس للنَّزعة المركزيَّة. في حين أنَّ القيمتين $\mathbf{Q}_1$ و $\mathbf{Q}_3$ ليستا من مقاييس النَّزعة المركزيَّة، ولكنّهما مقاييس موضع فقط.

٣- من استخدامات الرُّبَيْعيّات تحديد إن كانت قيمة $x$ من مجموعة بيانات مُعطاة هي قيمة متطرفة Extreme Value (**أو قيمة منعزلة Outlier Value**) أم لا. حيث يُقال عن قيمة $x$ من مجموعة بيانات مُعطاة إنَّها قيمة متطرفةً إذا حقَّقت إحدى العلاقتين (*) و (**) الآتيتين:

$$x < \mathbf{Q}_1 - 1.5(\mathbf{Q}_3 - \mathbf{Q}_1 \qquad (*)$$

وفي هذه الحالة نقول إنَّ القيمة $x$ متطرّفة بصغرها، وأمّا القيمة $\mathbf{Q}_1 - 1.5(\mathbf{Q}_3 - \mathbf{Q}_1)$ فإنَّها تُدعى Lowest Fence وسنرمز لها بـ **LF** أي إنَّه لدينا:

$$\mathbf{LF} \triangleq \mathbf{Q}_1 - 1.5(\mathbf{Q}_3 - \mathbf{Q}_1)$$

وهي تمثِّل في الواقع **الحد الأعلى للقيم المتطرفة بصغرها** لدى مجموعة بيانات مُعطاة.

$$x > \mathbf{Q}_3 + 1.5(\mathbf{Q}_3 - \mathbf{Q}_1) \qquad (**)$$

وفي هذه الحالة نقول إنَّ القيمة $x$ متطرّفة بكبرها، وأمّا القيمة $\mathbf{Q}_3 + 1.5(\mathbf{Q}_3 - \mathbf{Q}_1)$ فإنَّها تُدعى Highest Fence وسنرمز لها بـ **HF** أي إنَّه لدينا:

$$\mathbf{HF} \triangleq \mathbf{Q}_3 + 1.5(\mathbf{Q}_3 - \mathbf{Q}_1)$$

وهي تمثِّل في الواقع **الحد الأدنى للقيم المتطرفة بكبرها** لدى مجموعة بيانات مُعطاة.

فعلى سبيل المثال لو أخذنا البيانات الآتية:

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 6 | 7 | 7 | 4 | 15 | 5 | 5 | 4 | 5 | 3 | 4 | 5 | 7 | 3 | 4 |
| 1 | 4 | 5 | 4 | 4 | 7 | 5 | 5 | 6 | 5 | 7 | 1 | 6 | 4 | 7 |

فنجد أنَّ $\mathbf{Q}_1 = 4$ و $\mathbf{Q}_3 = 6.25$ ، ومن ثَّم ينتج لدينا $\mathbf{Q}_3 - \mathbf{Q}_1 = 6.25 - 4 = 2.25$ ، وبالتالي يكون:

$$\mathbf{HF} = \mathbf{Q}_3 + 1.5(\mathbf{Q}_3 - \mathbf{Q}_1) = 6.25 + (1.5)(2.25) = 9.625 < 15$$

$$\mathbf{LF} = \mathbf{Q}_1 - 1.5(\mathbf{Q}_3 - \mathbf{Q}_1) = 4 - (1.5)(2.25) = 0.625 < 1$$

فنجد أنَّ القيمة 15 متطرّفة بكبرها ولا وجود لقيمة متطرّفة بصغرها.

## (٢,٣) المئينات

## Percentiles

النوع الآخر من مقاييس الموضع هو ما يُعرف باسم "المئينات". في الواقع إنَّ المئين يوفّر معلومات حول كيفية امتداد البيانات من أصغر قيمة إلى أكبر قيمة، فعلى سبيل المثال يُذكر في كثير من الأحيان في جدول نتائج اختبار القبول للكلّيات والجامعات من حيث النسب المئوّية، فيقال على سبيل المثال إنَّ 65% فقط من المتقدمين استوفوا الشروط المطلوبة للتعيين. أو إنَّ 85% فقط من الطلاب حقّقوا درجة النجاح في المقرّر. لكن التعريف الرياضيّاتي لهذا المفهوم تقدِّمه الفقرة الآتية.

### (١ , ٣ , ٢) تعريف (المئيني الرائيّ $r$-Percentile):

المئينات هي تسعٌ وتسعون قيمة عددية تقسم البيانات بعد ترتيبها تصاعدياً إلى مئة شريحة (أو جزء) متساوية الأعداد من القيم، بحيث يكون في كل شريحة العدد نفسه من قيم البيانات. فلو كانت لدينا $x_1$ ، $x_2$ ،... و $x_n$ مجموعة من البيانات (عيِّنة أو لمجتمع الفارق هو استخدام الدليل $n$ أو $N$ فقط) المرتّبة تصاعدياً، فعندئذ القيمة التي يقع قبلها $r$ % من البيانات (من البيانات المرتّبة) وبعدها $(100-r)$ % تُدعى المئيني الرائيّ r-Percentile (r[th] Percentile أو المئيني ذو الرقم $r$)، وسنرمز له بـ $\mathbf{P}_r$ .

من تعريف المئينات نلاحظ أنَّ $\mathbf{P}_{25}=\mathbf{Q}_1$ ، $\mathbf{P}_{50}=\mathbf{Q}_2=\tilde{x}$ وكذلك $\mathbf{P}_{75}=\mathbf{Q}_3$ . كذلك يُلاحظ بسبب العدد الكبير للمئينات أنَّ هذا النوع من المقاييس يكون أكثر وضوحاً كلما ازداد عدد البيانات، ولذلك تكون نتائج هذا النوع من المقاييس أكثر دقَّة عندما يكون عدد البيانات كبيراً.

### (٢ , ٣ , ٢) تعيين المئينات

توجد طرائق عديدة لتعيين المئينات، وتتفاوت نتائجها وفقاً لهذه الطريقة أو تلك. في كتابنا هذا سنقدّم إحدى الطرائق المستخدمة في البرامج الإحصائية التي تعدّ من أكثر الطرائق دقةً.

لتعيين المئينات سنميِّز بين حالتين:

**١– من أجل البيانات الخام:**

لتكن $x_1$ ، $x_2$ ،... و $x_n$ بيانات خام مرتّبة تصاعدياً (عيِّنة أو مجتمع، والفارق هو استخدام الدليل $n$ أو $N$ فقط)، ولنضع بالتعريف:

$$p_r := \frac{r\,(n+1)}{100} \quad ; r \in \mathbb{N}_{99} \qquad [2,21]$$

الذي يُدعى رُتْبة المئين $\mathbf{P}_r$ في البيانات المرتّبة (وهو يحدِّد موضع المئين بين البيانات المرتّبة). عندئذ لتعيين قيمة المئين $\mathbf{P}_r$ من أجل أي $r \in \mathbb{N}_{99}$ نقوم بما يلي:

بفرض أنَّ $k$ هو الجزء الصحيح من المقدار $p_r$ ، وأنَّ الباقي منه يساوي $s$ ، فعندئذ تُحسب قيمة المئين $\mathbf{P}_r$ من خلال العلاقة الآتية:

$$\mathbf{P}_r = x_k + s\,(x_{k+1} - x_k) \qquad ; r \in \mathbb{N}_{99} \qquad [2,22]$$

نلاحظ هنا أنّه إذا كان $s=0$ أو $x_{k+1}=x_k$ فإنَّه سيكون لدينا $\mathbf{P}_r = x_k$ فقط.

**٢– في حالة البيانات المجمَّعة:**

بفرض أنَّه لدينا بيانات مُجمَّعة في جدول توزيع تكراري كما في الجدول (١ , ٢) ولنضع بالتعريف:

$$p_r := \frac{r \cdot \sum f_i}{100} \quad ; r \in \mathbb{N}_{99} \qquad [2,23]$$

عندئذ يُعيَّن المئين ذا الرقم $r$ لبيانات هذا الجدول من خلال العلاقة الآتية:

$$\mathbf{P}_r := L_{P_r} + \frac{p_r - \left(F_{P_r} - f_{P_r}\right)}{f_{P_r}} \times \mathbf{C} \qquad ; r \in \mathbb{N}_{99} \qquad [2,24]$$

علماً أنَّ $L_{P_r}$ هو الحد الأدنى لفئة المئين $\mathbf{P}_r$ ، وأنَّ الفئة الخاصّة بالمئين $\mathbf{P}_r$ هي أول فئة تكرارها المتجمِّع الصاعد أكبر أو يساوي $p_r$ ، $f_{P_r}$ هو تكرار الفئة الخاصّة بالمئين $\mathbf{P}_r$ ، و $F_{P_r}$ هو التكرار المتجمّع الصاعد للفئة الخاصّة بالمئين $\mathbf{P}_r$ ، وأخيراً $\mathbf{C}$ هي سعة الفئة الخاصّة بالمئين $\mathbf{P}_r$ (هنا لم نرفق دليلًا لـ $\mathbf{C}$ لأنَّ جميع الفئات متساوية السعة).

في حال كانت البيانات مقدَّمة من خلال جدول تكراري، فعندئذ نعيد تنظيم الجدول بحيث تصبح قيم الممثِّلين في الجدول التكراري (٢,١) مرتّبةً تصاعدياً، وبعد ذلك نفترض أنَّ البيانات الأصل هي $y_1$ ، $y_2$، ... و $y_n$ مع $n = \sum f_i$ و $y_1 < y_2 < ... < y_n$ ، ومن ثَمَّ تُحسب رتبة المئين $r$ بوساطة العلاقة [2,21] لهذه البيانات، فإن كانت قيمة $p_r$ عدداً صحيحاً ووجدت قيمة ممثِّل $x_i$ في الجدول التكراري لها تكرارها تراكمي يساوي $p_r$ فإنَّنا نأخذ تلك القيمة كقيمة للمئين $r$ ، وخلاف ذلك نأخذ أول قيمة لممثِّل $x_i$ في الجـــدول التكراري لها تكرارها تراكمي أكبر من $p_r$ كقيمة للمئين $r$ . أمّــا إذا كانت قيمة $p_r$ عدداً غير صحيح، فعندئذ نطبِّق على البيانات المرتَّبة $y_1$ ، $y_2$، ... و $y_n$ العلاقة الآتية:

$$\mathbf{Q}_r = y_\ell + s\,(y_{\ell+1} - y_\ell) \;\; ; r \in \mathbb{N}_{99}$$

علماً أنَّ $\ell$ هنا هو الجزء الصحيح من $p_r$ و $y_\ell$ هي أول قيمة ممثِّل $x_i$ في الجدول التكراري لها تكرارها تراكمي أكبر أو يساوي $\ell$ ، و $y_{\ell+1}$ هي القيمة $x_i$ نفسها إذا كان $F_i > \ell$ ، وخلاف ذلك (أي أنَّ $F_i = \ell$ ) تكون $y_{\ell+1}$ هي $x_{i+1}$ من الجدول التكراري.

### (٢,٣,٢,١) أمثلة

١- بالعودة إلى المثال / ١/ من (٢,١,٢,١) وبعد ترتيب تلك البيانات تصاعدياً يصبح لها العرض الآتي:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| 147 | 160 | 168 | 171 | 172 | 173 | 177 | 180 | 189 |
| 149 | 160 | 168 | 171 | 172 | 173 | 178 | 181 | 190 |
| 149 | 160 | 169 | 171 | 172 | 173 | 178 | 181 | 190 |
| 152 | 163 | 169 | 171 | 172 | 173 | 178 | 181 | 192 |
| 147 | 160 | 168 | 170 | 172 | 173 | 177 | 180 | 189 |
| 152 | 163 | 169 | 171 | 173 | 173 | 178 | 182 | 192 |
| 154 | 163 | 169 | 171 | 173 | 175 | 178 | 182 | 193 |
| 155 | 163 | 169 | 171 | 173 | 175 | 178 | 183 | 193 |
| 157 | 163 | 169 | 171 | 173 | 175 | 179 | 183 | 193 |
| 157 | 165 | 169 | 172 | 173 | 175 | 179 | 184 | 195 |
| 158 | 167 | 169 | 172 | 173 | 175 | 179 | 185 | 195 |
| 158 | 167 | 169 | 172 | 173 | 175 | 179 | 185 | 196 |
| 158 | 167 | 170 | 172 | 173 | 175 | 179 | 188 | |
| 160 | 167 | 170 | 172 | 173 | 175 | 179 | 188 | |
| 160 | 168 | 170 | 172 | 173 | 175 | 180 | 189 | |

بيانات المثال / ١/ من (٢,١,٢,١) بعد ترتيب تصاعدياً

ولنقم بتعيين المئين $\mathbf{P}_{35}$ و $\mathbf{P}_{89}$ لهذه البيانات.

نعلم أنَّ عدد البيانات $n = 132$ ، ومن ثَمَّ تكون رتبة المئين $\mathbf{P}_{35}$ و $\mathbf{P}_{89}$ على الترتيب هي:

$$p_{35} = \frac{35(n+1)}{100} = \frac{35(132+1)}{100} = 46.90 \qquad \& \qquad p_{89} = \frac{89(n+1)}{100} = \frac{89(132+1)}{100} = 119.26$$

ومن ثَمَّ يكون لدينا:

$$\mathbf{P}_{35} = x_{46} + 0.90(x_{47} - x_{46}) = 170 + 0.90(171 - 170) = 170.90$$
$$\mathbf{P}_{89} = x_{119} + 0.26(x_{120} - x_{119}) = 189 + 0.40(189 - 189) = 189$$

٢- لنعد إلى بيانات المثال / ١/ من (١-٢-٢-٢) حيث لدينا $n = 20$ ، ولنقم بحساب المئينين $\mathbf{P}_{15}$ و $\mathbf{P}_{85}$ .

العمل: من أجل ذلك نجد رُتْبَة المئين $\mathbf{P}_{15}$ و $\mathbf{P}_{85}$ على الترتيب هي:

$$p_{15} = \frac{15(n+1)}{100} = \frac{15(20+1)}{100} = 3.15 \qquad \& \qquad p_{85} = \frac{85(n+1)}{100} = \frac{85(20+1)}{100} = 17.85$$

ومن ثَمَّ يكون لدينا:

$$\mathbf{P}_{15} = x_3 + 0.15\,(x_4 - x_3) = 483 + 0.15(485 - 483) = 483 + 0.30 = 483.30$$
$$\mathbf{P}_{85} = x_{17} + 0.85(x_{18} - x_{17}) = 535 + 0.85(537 - 535) = 535 + 1.7 = 536.7$$

٣- بالعودة إلى بيانات المثال / ٢/ من (١ , ٢ , ٢ , ٢) المُعطاة من خلال الجدول (٣ , ٢)، ولنقم بتعيين قيم المئينات $\mathbf{P}_{25}$ و $\mathbf{P}_{50}$ و $\mathbf{P}_{75}$ لبيانات ذلك الجدول.

من أجل ذلك نلاحظ أنَّ رتب هذه المئينات على الترتيب هي:

$$p_{25} = \frac{25\sum_{i=1}^{k} f_i}{100} = \frac{25 \times 48}{100} = 12 \quad \& \quad p_{50} = \frac{50\sum_{i=1}^{k} f_i}{100} = 24 \quad \& \quad p_{75} = \frac{75\sum_{i=1}^{k} f_i}{100} = 36$$

وأنَّ الفئة الخاصّة بالمئين $\mathbf{P}_{25}$ هي الفئة الثانية لأنَّها أول فئة تكرارها المتجمِّع الصّاعد أكبر أو يساوي $p_{25}$ ، ومن ثَمَّ يكون لدينا:

$$\mathbf{P}_{25} = L_{P_{25}} + \frac{p_{25} - \left(F_{P_{25}} - f_{P_{25}}\right)}{f_{P_{25}}} \times \mathbf{C} = 30 + \frac{12 - (14 - 19)}{9} \times 5 = 33.89$$

وكذلك الفئة الخاصّة بالمئين $\mathbf{P}_{50}$ هي الفئة الثالثة لأنَّها أول فئة تكرارها المتجمِّع الصّاعد أكبر أو يساوي $p_{50}$ ، ومن ثَمَّ يكون لدينا:

$$\mathbf{P}_{50} = L_{P_{50}} + \frac{p_{50} - \left(F_{P_{50}} - f_{P_{50}}\right)}{f_{P_{50}}} \times \mathbf{C} = 35 + \frac{24 - (14 - 19)}{9} \times 5 = 45.56$$

وأخيراً الفئة الخاصّة بالمئين $\mathbf{P}_{75}$ هي الفئة الرابعة لأنَّها أول فئة تكرارها المتجمِّع الصّاعد أكبر أو يساوي $p_{75}$ ، ومن ثَمَّ يكون لدينا:

$$\mathbf{P}_{75} = L_{P_{75}} + \frac{p_{75} - \left(F_{P_{75}} - f_{P_{75}}\right)}{f_{P_{75}}} \times \mathbf{C} = 40 + \frac{36 - (14 - 19)}{9} \times 5 = 57.22$$

لاحظ أنَّ هذه المئينات التي تمَّ حسابها هي الرُّبَيْعيّات الثلاثة $\mathbf{Q}_1$ و $\mathbf{Q}_2$ و $\mathbf{Q}_3$ .

٣- بالرجوع إلى المثال /٣/ من (١, ٦, ١, ٢). لنقم بتعيين المئين الثلاثة $\mathbf{P}_{35}$ لمجموعة البيانات $\mathbf{Z}$، فنجد أنَّ رتبة المئين $\mathbf{P}_{35}$ هي:

$$p_{35} = \frac{35(n+1)}{100} = \frac{35(30+1)}{100} = 10.85$$

ومن ثمَّ الجزء الصحيح من $p_{35}$ هو $\ell = 10$، ومنه تكون قيمة $y_\ell = z_3 = 11$، وبما أنَّ $F_3 = 15 > \ell = 10$ فإنَّ $y_{\ell+1} = z_3 = 11$ أيضاً، وبالتالي يكون لدينا:

$$\mathbf{Q}_1 = y_{10} + 0.85(y_{11} - y_{10}) = z_3 + 0.85(z_3 - z_3) = z_3 = 11$$

### (٢, ٣, ٢, ٢) ملاحظات

١- نشير هنا إلى أنَّ المئينات تستخدم عادة عندما يكون حجم البيانات كبيراً بقدر كاف (وغالباً أكثر من 99 قيمة)، وأنَّ ما قُدِّم سابقاً وسيقدم لاحقاً من أمثلة على بيانات قليلة العدد كان من باب التوضيح والتبسيط فقط.

٢- توجد أكثر من طريقة لتعيين المئين الموافق لأيّة قيمة من قيم بيانات خام مُعطاة، ومنها الطريقتين الآتيتين:

**طريقة أولى:** لتكن $x_1$، $x_2$، ...و $x_n$ بيانات خام مُرتّبة تصاعدياً (لعيّنة أو لمجتمع الفارق هو استخدام الرمز $n$ أو $N$ فقط)، ومن أجل $1 \le i \le n$ لنرمز بـ $\mathbf{P}_{x_i}$ لقيمة المئين الموافقة للقيمة $x_i$، فعندئذ من أجل أيّة قيمة $i$ و $\mathbb{N}_n$ تُعطى قيمة $\mathbf{P}_{x_i}$ بوساطة العلاقة الآتية:

$$\mathbf{P}_{x_i} := \frac{i}{n+1} \times 100\ \% \quad \text{(تقرأ كنسبة مئوية)} \qquad [2,25\text{-a}]$$

**طريقة ثانية:** هذه الطريقة أقلّ دقّةً من الطريقة السابقة ولكنّها مقبولة لتعيين المئين الموافق لأيّة قيمة من قيم بيانات خام $x_1$، $x_2$، ...و $x_n$ بيانات خام مُرتّبة تصاعدياً، عندئذ تُعطى قيمة $\mathbf{P}_{x_i}$ من أجل أيّة قيمة $i$ و $\mathbb{N}_n$ بوساطة العلاقة الآتية:

$$\mathbf{P}_{x_i} := \frac{N_i + 0.5}{n} \times 100\ \% \quad \text{(تقرأ كنسبة مئوية)} \qquad [2,25\text{-b}]$$

علماً أنَّ $N_i$ هو عدد البيانات المرتَّبة التي أصغر من القيمة $x_i$.

### (٢, ٣, ٢, ٣) أمثلة:

١- لتكن لدينا البيانات الآتية:

0 -1 2 3 6 9 2

ولنستخدم العلاقة [2,25-a] لتعيين المئينات للبيانات المُعطاة.

بعد ترتيب هذه البيانات تصاعدياً نجد المئين الموافق لكل قيمة من القيم المرتبة كما في الجدول الآتي (قمنا باستخدام الجدول على سبيل التوضيح والإيجاز فقط):

الجدول (٢, ٥)

| $x_i$ | -1 | 0 | 2 | 2 | 3 | 6 | 9 | قيمة ما $x$ أكبر تماماً من $x_7$ |
|---|---|---|---|---|---|---|---|---|
| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $< x$ |
| $\mathbf{P}_{x_i}$ | 12.5 | 25 | 36.5 | 50 | 62.5 | 75 | 87.5 | 100 |

نلاحظ هنا أنّه لو طلب منّا حساب المئين $\mathbf{P}_{88}$ للبيانات المُعطاة فإنّه لن نستطيع تعيينه لأنَّ أكبر قيمة في البيانات مئنها 87.5.

٢- لتكن لدينا البيانات الآتية:

5 3 1 7 6 6 11 25 9 12

باستخدام العلاقة [2,25-a] لتعيين المئينات للبيانات المُعطاة نجدها كما في الجدول الآتي:

الجدول (٢,٦)

| $x_i$ | 1 | 3 | 5 | 6 | 6 | 7 | 9 | 11 | 12 | 25 | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $x_9$ | $x_{10}$ | $< x$ |
| $\mathbf{P}_{x_i}$ | 9.09 | 18.18 | 27.27 | 36.36 | 45.45 | 54.55 | 63.64 | 72.73 | 81.82 | 90.91 | 100 |

نلاحظ هنا أنَّه كلما زاد عدد البيانات فإنَّه يمكننا حساب مئينات من مراتب أعلى، وعلاوةً على ذلك يُلاحظ أنَّ القيمة المئينية لأية قيمة مُرتَّبة $x_i$ هي من مضاعفات القيمة المئينية لأصغر قيمة في البيانات $x_1$ ، وأنَّه لدى أية قيمة $x$ تالية لأكبر قيمة في البيانات سيكون لدينا $\mathbf{P}_{x_i} = 100\%$ ليشير بذلك على أنَّ جميع البيانات تقع قبل هذه القيمة $x$ .

٣- لتكن لدينا البيانات الآتية:

9 -3 5 -2 9 7 6 4 4 6 3

ولنستخدم العلاقة [2,25-b] لتعيين المئينات للبيانات المُعطاة. بالطبع بعد ترتيب هذه البيانات تصاعدياً فإنَّ المئين الموافق لكل قيمة من هذه القيم المرتّبة نجدها تساوي القيمة المقابلة لها في الجدول الآتي:

الجدول (٢,٧)

| $x_i$ | -3 | -2 | 3 | 4 | 4 | 5 | 6 | 6 | 7 | 9 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | $x_1$ | $x_2$ | $x_3$ | $x_4$ | $x_5$ | $x_6$ | $x_7$ | $x_8$ | $x_9$ | $x_{10}$ | $x_{11}$ |
| $N_i$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| $\mathbf{P}_{x_i}$ | 4.5 | 13.6 | 22.7 | 31.8 | 40.9 | 50 | 59.1 | 68.2 | 77.3 | 86.4 | 95.5 |

فلو أردنا على سبيل المثال تعيين المئين $\mathbf{P}_{50}$ للبيانات المُعطاة ،فإنَّنا نجد ما يلي:

$$p_{50} = \frac{50(n+1)}{100} = \frac{50(11+1)}{100} = 6.00 \quad \Rightarrow \quad \mathbf{P}_{50} = x_6 = 5$$

فنجد أنَّها تتوافق مع نتيجة الطريقة التي تستخدم العلاقة [2,25-b]، حيث كان للقيمة $x_6 = 5$ المئين $\mathbf{P}_{x_6} = 50$ ، أي إن $\mathbf{P}_{x_6} = \mathbf{P}_{50}$ ، ولكن كما ذكرنا سابقاً فإنَّ هذه الطريقة ليست دقيقة، فمن الممكن أن تكون قيمة $\mathbf{P}_{x_i}$ أكبر أو أصغر من القيمة الحقيقية للمئين، فعلى سبيل المثال نجد من أجل $\mathbf{P}_{25}$ و $\mathbf{P}_{75}$ ما يلي:

$$p_{25} = \frac{25(n+1)}{100} = \frac{25(11+1)}{100} = 3.00 \;\Rightarrow\; \mathbf{P}_{25} = x_3 = 3 \;\Rightarrow\; \mathbf{P}_{x_3} < \mathbf{P}_{25}$$

$$p_{75} = \frac{75(n+1)}{100} = \frac{75(11+1)}{100} = 9.00 \;\Rightarrow\; \mathbf{P}_{75} = x_9 = 7 \;\Rightarrow\; \mathbf{P}_{x_9} > \mathbf{P}_{75}$$

## (٢,٤) مقاييس التشتت

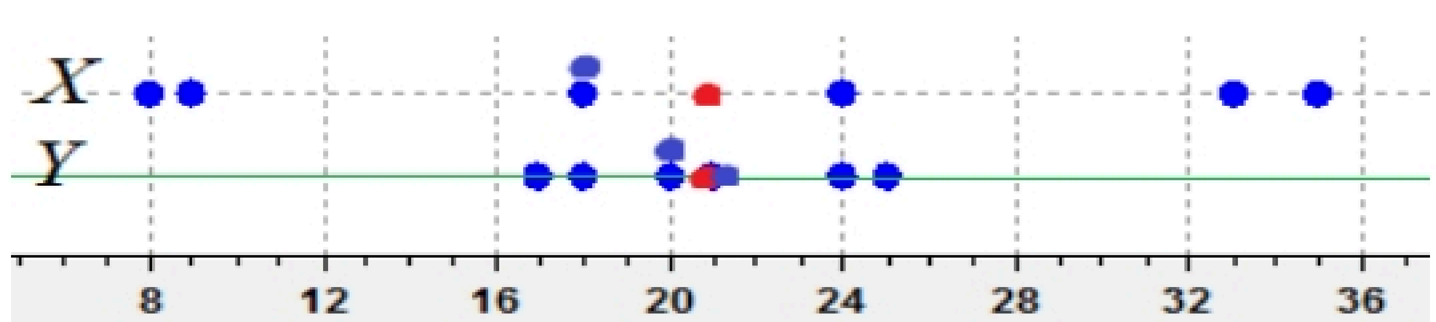
## Dispersion Measures

لقد قدمنا في الفقرة السّابقة عرضاً لأهم مقاييس النَّزعة المركزيَّة وأخرى للموضع، وكان لكل منها ميزاته التي تجعل منه مقياساً مفضلاً من أجل حالات مُحدَّدة، ولكن هذه المقاييس للنَّزعة المركزيَّة لا تقدِّم وصفاً دقيقاً لتوزيع البيانات، ذلك أنَّه من المهم معرفة كيفية توزّع هذه البيانات حول القيمة الممثِّلة لمقياس نزعتها المركزيَّة (متوسطها، وسيطها أو منوالها) أيضاً، حيث يسأل المرء إن كانت البيانات قريبة من هذه القيمة أم أنَّها متناثرة بشكل متباعد عنها، فعلى سبيل المثال لو أخذنا مجموعتي البيانات الآتية حيث تمثِّل كل منها درجات الحرارة في مدينة مأخوذةً كل ساعتين خلال يوم كامل:

الجدول (٨ , ٢)

| الوقت | 6 | 8 | 10 | 12 | 14 | 16 | 18 |
|---|---|---|---|---|---|---|---|
| المدينة $X$ | 8 | 9 | 18 | 35 | 33 | 24 | 18 |
| المدينة $Y$ | 17 | 18 | 20 | 25 | 24 | 21 | 20 |

فنجد أنَّ متوسط درجات الحرارة في كلا المدينتين خلال هذا اليوم يساوي 20.71 إلا أنَّ تبعثر قيم درجات الحرارة حول متوسطها يختلف من مدينة لأخرى، حيث نلاحظ أنَّ قيم درجات الحرارة في المدينة $X$ لها توضّعات مختلفة تماماً عما هو عليه الحال لدى قيم درجات الحرارة في المدينة $Y$ (انظر الشكل الآتي).



الشكل (٦ , ٢).

إذن لابدَّ من وجـود مقاييس تُوضِّح لنا كيفية تبعثر (أو تشتّت) البيانات حول مقياس نزعتها المركزيَّة لكي يتكوَّن لدينا انطباع أكثر وضوحاً حول سلوك البيانات.

سنقدّم فيما يلي أهم مقاييس التشتّت (أو مقاييس التبعثر **Scatter Measures** أو مقاييس الانتشار **Spread Measures**) التي تَندرج تحت مُسَمَّى **مقاييس الاختلاف** Variability Measures.

في الواقع لوحظ أنَّه يمكن النظر للفروقات بين قيم البيانات عن متوسطها كمقياس للتشتّت، ولكن نعلم أنَّ المجموع الجبري لهذه الفروق يساوي الصفر دوماً، لذلك من الممكن أخذ القيم المطلقة للفروق بين البيانات ومتوسطها من أجل استخدامها كمقياس للتشتّت، وعلى هذا الأساس بُني تعريف **الانحراف المتوسط** Mean Deviation (أو مُعدَّل الانحرافات **Average Deviations**) حيث يُستخدم في حسابه كل قيم البيانات. في الواقع يقوم الانحراف المتوسط على أخذ المتوسط للمسافات بين كل قيمة والمتوسط لهذه البيانات، ومن ثَمَّ يُعطينا قراراً معقولاً ومقبولاً حول كيفية تبعثر البيانات حول متوسطها.

### (١ , ٤ , ٢) الانحراف المتوسط Mean Deviation

من أجل تعريف **الانحراف المتوسط** لمجموعة بيانات (ويُرمز له بـ **MD**)، سنناقش الحالتين الآتيتين:

**أولاً: من أجل البيانات الخام:**

أ- إذا كانت $x_1$، $x_2$، ...، و $x_n$ بيانات عيِّنة بمتوسط $\overline{x}$ ، فعندئذٍ يُعرَّف الانحراف المتوسط لهذه البيانات بوساطة العلاقة الآتية:

$$\mathbf{MD} := \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \overline{x}\right| \qquad [2,26\text{-a}]$$

ب- إذا كانت $x_1$، $x_2$، ...و $x_n$ بيانات مجتمع بمتوسط $\mu$ ، فعندئذ يُعرَّف الانحراف المتوسط لهذه البيانات بوساطة العلاقة الآتية:

$$\mathbf{MD} := \frac{1}{N}\sum_{i=1}^{N}\left|x_i - \mu\right| \qquad [2,26\text{-b}]$$

كما ذكرنا سابقاً فإنَّ الانحراف المتوسط لمجموعة بيانات هو المتوسط الحسابي للمسافات الفاصلة بين قيم البيانات ومتوسطها.

**ثانياً: من أجل البيانات المُجمَّعة:**

بفرض أنَّه لدينا بيانات مُجمَّعة في جدول توزيع تكراري كما في الجدول (١ , ٢)، فعندئذ:

أ- إذا كانت هذه البيانات لعيِّنة متوسطها $\overline{x}$ ، فعندئذ يُعرَّف الانحراف المتوسط لهذه البيانات من خلال العلاقة الآتية:

$$\mathbf{MD} := \frac{1}{\sum_{i=1}^{k} f_i}\sum_{i=1}^{k} f_i \cdot \left|x_i - \overline{x}\right| \qquad [2,27\text{-a}]$$

ب- إذا كانت هذه البيانات لمجتمع متوسطه $\mu$ ، فعندئذ يُعرَّف الانحراف المتوسط لهذه البيانات من خلال العلاقة الآتية:

$$\mathbf{MD} := \frac{1}{\sum_{i=1}^{k} f_i}\sum_{i=1}^{k} f_i \cdot \left|x_i - \mu\right| \qquad [2\text{–}27\text{-b}]$$

**(١ , ١ , ٤ , ٢) أمثلة**

١ – لتكن لدينا البيانات الآتية التي تُمثِّل الأجور الساعية لعينتين من العمال في مهنتين مختلفتين (مُقدَّرة بوحدة نقدية ما):

الجدول (٩ , ٢)

| العامل | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| الأجر الساعي للمهنة $X$ | 3 | 8 | 5 | 10 | 7 | 15 | 7 | 9 |
| الأجر الساعي للمهنة $Y$ | 7 | 8 | 5 | 8 | 17 | 6 | 8 | 5 |

فنجد أنَّ قيم المتوسط لمجموعتي البيانات $X$ و$Y$ هما $\overline{x} = 8$ و $\overline{y} = 8$ ، وأمّا الانحراف المتوسط لكل منهما فإنَّه يساوي:

$$\mathbf{MD_X} = \frac{1}{8}\sum_{i=1}^{8}\left|x_i - \overline{x}\right| = 2.50 \qquad \& \qquad \mathbf{MD_Y} = \frac{1}{8}\sum_{i=1}^{8}\left|y_i - \overline{y}\right| = 2.25$$

وهكذا يُظهر لنا هذا المقياس أنَّ تشتّت البيانات $X$ حول متوسطها أكبر من تشتّت البيانات $Y$ حول متوسطها.

٢- بالرجوع إلى جدول التوزيع التّكراري (٢ , ٢) في المثال / ٢/ من (١ , ٢ , ١ , ٢) نجد أنَّ قيمة الانحراف المتوسط لتلك البيانات يساوي:

$$\mathbf{MD} = \frac{1}{\sum_{i=1}^{8} f_i}\sum_{i=1}^{8} f_i \left|x_i - \overline{x}\right| = \frac{461.2}{50} = 9.224$$

**(٢ , ١ , ٤ , ٢) ملاحظة**

من الممكن حساب الانحراف المتوسط حول الوسيط (وسنرمز له بـ $\mathbf{M\tilde{D}}$ ) أو حول المنوال (في حال وجوده ووحدانيته، وسنرمز له بـ $\mathbf{M\hat{D}}$ ) وذلك على النحو الآتي (حيث تُستخدم هذه المقاييس في ضبط الجودة الإحصائي وبعض الدراسات الاجتماعية).

**أولاً: من أجل البيانات الخام:**

$$\mathbf{M\tilde{D}} := \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \tilde{x}\right| \quad \text{for a sample with median } \tilde{x}$$

$$\mathbf{M\tilde{D}} := \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \tilde{\mu}\right| \quad \text{for a population with median } \tilde{\mu}$$

$$\mathbf{M\hat{D}} := \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \hat{x}\right| \quad \text{for a sample with mode } \hat{x}$$

$$\mathbf{M\hat{D}} := \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \hat{\mu}\right| \quad \text{for a population with mode } \hat{\mu}$$

**ثانياً: من أجل البيانات المجمَّعة:**

$$\mathbf{M\tilde{D}} := \frac{1}{\sum_{i=1}^{k} f_i}\sum_{i=1}^{k} f_i \cdot \left|x_i - \tilde{x}\right| \quad \text{for a sample with median } \tilde{x}$$

$$\mathbf{M\tilde{D}} := \frac{1}{\sum_{i=1}^{k} f_i}\sum_{i=1}^{k} f_i \cdot \left|x_i - \tilde{\mu}\right| \quad \text{for a population with median } \tilde{\mu}$$

$$\mathbf{M\hat{D}} := \frac{1}{\sum_{i=1}^{k} f_i}\sum_{i=1}^{k} f_i \cdot \left|x_i - \hat{x}\right| \quad \text{for a sample with mode } \hat{x}$$

$$\mathbf{M\hat{D}} := \frac{1}{\sum_{i=1}^{k} f_i}\sum_{i=1}^{k} f_i \cdot \left|x_i - \hat{\mu}\right| \quad \text{for a population with mode } \hat{\mu}$$

**(٣, ١, ٤, ٢) مثال**

لتكن لدينا البيانات الآتية التي تمثِّل عدد الأخطاء المرتكبة من قبل لاعبي كرة القدم لأحد الأندية في مباريات للدوري العام.

الجدول (٢, ١٠)

| اللعبة | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| عدد الأخطاء | 7 | 5 | 7 | 6 | 2 | 9 | 4 | 5 | 7 | 3 |

فنجد أنَّ $\overline{x} = 5.5$ و $\tilde{x} = 5.5$ و $\hat{x} = 7$ ، ومن يكون لدينا:

$$\mathbf{MD} = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \overline{x}\right| = \frac{1}{10}\sum_{i=1}^{10}\left|x_i - \overline{x}\right| = \frac{19}{10} = 1.9$$

$$\mathbf{M\tilde{D}} = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \tilde{x}\right| = \frac{1}{10}\sum_{i=1}^{10}\left|x_i - \tilde{x}\right| = \frac{19}{10} = 1.9$$

$$\mathbf{M\hat{D}} = \frac{1}{n}\sum_{i=1}^{n}\left|x_i - \hat{x}\right| = \frac{1}{10}\sum_{i=1}^{10}\left|x_i - \hat{x}\right| = \frac{17}{10} = 1.7$$

إذا قمنا بمقارنة هذه النتائج مع ما حصلنا عليه حول المتوسط نجد أنَّ $\mathbf{MD} = \mathbf{M\tilde{D}} \neq \mathbf{M\hat{D}}$ .

لقد لاحظنا فيما سبق أن المتوسط الحسابي للقيم المطلقة لفروقات قيم البيانات عن متوسطها يعطي مقياساً مقبولاً لمدى تبعثر البيانات حول متوسطها، ولكن وجد أنَّ هذاالمقياس يصبح أفضل إذا أُخذَ الجذر التربيعي الموجب للمتوسط الحسابي لمربعات الفروق بين البيانات ومتوسطها. إنَّ التعريف الآتي يمهِّد لهذا المقياس.

### (٢ , ٤ , ٢) تعريف (التباين Variance)

من أجل تعريف التباين لبيانات مُعطاة سنميِّز بين الحالتين الآتيتين:

**أولاً: من أجل البيانات الخام:**

أ ) إذا كانت البيانات $x_1$، $x_2$، ...و $x_n$ لعيِّنة متوسطها $\overline{x}$ ، فعندئذ يُعرَّف التباين لهذه البيانات (ويُرمز له بـ $S^2$) من خلال العلاقة الآتية:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2 \qquad [2,28\text{-a}]$$

ب) إذا كانت هذه البيانات هي $x_1$، $x_2$، ...و $x_N$ لمجتمع إحصائي متوسطه $\mu$ ، فعندئذ يُعرَّف التباين لهذه البيانات (ويُرمز له بـ $\sigma^2$) من خلال العلاقة الآتية:

$$\sigma^2 := \frac{1}{N} \sum_{i=1}^{N} \left(x_i - \mu\right)^2 \qquad [2,28\text{-b}]$$

**ثانياً: من أجل البيانات المجمَّعة:**

بفرض أنَّه لدينا بيانات مُجمَّعة في جدول توزيع تكراري كما في الجدول (١ , ٢)، فإذا كانت هذه البيانات:

أ - هي بيانات لعيِّنة حجمها $n = \sum_{i=1}^{k} f_i$ ومتوسطها $\overline{x}$ ، فعندئذ يُعرَّف التباين (ويُرمز له بـ $S^2$ أيضاً) لهذه البيانات من خلال العلاقة الآتية:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{k} f_i \left(x_i - \overline{x}\right)^2 \qquad [2,29\text{-a}]$$

ب- هي بيانات مجتمع حجمه $N = \sum_{i=1}^{k} f_i$ ومتوسطه $\mu$ ، فعندئذ يُعرَّف التباين (ويُرمز له بـ $\sigma^2$ أيضاً) لهذه البيانات من خلال العلاقة الآتية:

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^{k} f_i \left(x_i - \mu\right)^2 \qquad [2,29\text{-b}]$$

### (١ , ٢ , ٤ , ٢) ملاحظات

١- إنَّ استخدام رمز التربيع فوق الرمز $S$ أو $\sigma$ للدلالة على أنَّ التباين هو مقدار غير سالب، وأنَّ القيمة النّاتجة عنه تقرأ بالوحدة المربّعة لوحدة القياس المستخدمة في البيانات، فعلى سبيل المثال إذا كانت وحدة القياس للبيانات هي المتر فإنَّ قيمة التباين لهذه البيانات تُقرأ بالمتر المربّع، ولهذا السبب فإنَّ قيمة التباين لا تستخدم بشكلٍ مباشر كمقياس للتشتت.

٢- إنَّ الصيغة العامّة لتباين مجموعة من القيم العشوائية والمتساوية النصيب في الاختيار $x_1$، $x_2$، ...و $x_n$ وذات متوسط $\overline{x}$ هي (بشكلها الرياضياتي - وفقاً لنظرية الاحتمالات):

$$\tilde{S}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2 \qquad [2,30]$$

التي تمثِّل التباين لمتغيِّر عشوائي خاضع للتوزيع المنتظم المتقطِّع بوسيط $n$ (انظر الفصلين السادس والسابع)، ولهذا فإنَّ صيغة التباين للمجتمع الإحصائي (الذي يمثِّلها $\sigma^2$) توافق الصيغة العامّة للتباين. أمّا إذا كانت هذه البيانات هي بيانات عيِّنة بحجم $n$ ، فإنَّ صيغة التباين تُكتب بأحد الشكلين الآتيين:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2 \qquad \text{for raw data}$$

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{k} f_i\left(x_i - \overline{x}\right)^2 \qquad \text{for grouped data}$$

وكلا هاتين العلاقتين تُعرف باسم **التباين العملي** (أو تباين العيِّنة) Empirical Variance للبيانات، وتبرير استخدامه من أجل العينات تجده في الفصل العاشر. لكن عندما يكون حجم العينية $n$ كبيراً (أي إنّ $n \le 32$)، فإنَّه يُنظر إلى الفارق بين التباين والتباين العملي على أنّه مقدار صغير يمكن إهماله، وإذا كان حجم العينية صغيراً (أي إنّ $n < 32$)، فإنَّه يُفضَّل استخدام صيغة التباين العملي. أمّا من أجل البيانات المُجمّعة في جداول فإنَّه غالباً ما يكون عددها كبيراً، ومن ثمَّ استخدام صيغة التباين العملي ستعطي قيمة قريبة جداً من القيمة النّاتجة عن الصيغة العامّة للتباين، ولذلك استخدام أي الصيغتين سيكون مقبولاً، ولكن في حال كان مجموع التكرارات صغيراً فإنَّه يجب استخدام صيغة التباين العملي.

٣- إنَّ مجموع مربّعات الانحرافات لبيانات $x_1$، $x_2$،...،و $x_n$ عن أي عدد حقيقي $a$ يكون أصغرياً عندما يكون $a=\overline{x}$. إذ إنَّه باستخدام مفهوم النهايات الحدّية (راجع مفهوم النهايات الحدّية في حساب التفاضل) يكون لدينا:

$$\left(\sum_{i=1}^{n}(x_i - a)^2\right)^{\backslash} = 2\sum_{i=1}^{n}\ (x_i - a) = 2\sum_{i=1}^{n}x_i - 2na = 0$$

والتي ينتج عنها أنَّ $2na = 2\sum_{i=1}^{n}x_i$، ومن ثمَّ يكون لدينا $a = \frac{1}{n}\sum_{i=1}^{n}x_i$، ولهذا السبب يُنظر إلى مفهوم التباين على أنّه أفضل مقياس للتشتّت.

٤- إنَّ الجذر التربيعي الموجب للتباين يُعرف باسم **الانحراف المعّياري** Standard Deviation، وهو يَستخدم وحدة قياس البيانات نفسها، ويُرمز له بـ $S$ إذا كانت البيانات لعيِّنة، وأمّا إذا كانت البيانات لمجتمع فإنَّه يُرمز له بـ $\sigma$، أي إنَّه لدينا:

$$\sigma := +\sqrt{\sigma^2} \qquad \text{for a population} \qquad [2,31\text{-a}]$$

$$S := +\sqrt{S^2} \qquad \text{for a sample} \qquad [2,31\text{-b}]$$

ويُعرف الجذر التربيعي الموجب لتباين العيّنة (العلاقة [2,31-b]) باسم **الانحراف المعّياري العملي** Empirical Standard Deviation (أو الانحراف المعيّاري للعينة).

٥- إنَّ تباين مجموعة من البيانات يمكن أن يُحسب من خلال العلاقة الآتية أيضاً وهي تَستخدم قيم البيانات بشكل مباشر دون اللجوء إلى حساب المتوسط أو معرفة قيمته:

$$\sigma^2 = \frac{1}{N^2}\left[N\ \sum_{i=1}^{N}x_i^2 - \left(\sum_{i=1}^{N}x_i\right)^2\right] \qquad \text{for population} \qquad [2,32,a]$$

$$S^2 = \frac{1}{n(n-1)}\left[n\ \sum_{i=1}^{n}x_i^2 - \left(\sum_{i=1}^{n}x_i\right)^2\right] \qquad \text{for sample} \qquad [2,32,b]$$

**(٢ ,٢ ,٤ ,٢) أمثلة**

١- لتكن لدينا البيانات الآتية التي تمثِّل تكلفة إصلاح 60 جهاز حاسب (مقدَّرةً بالدولار) في مُعامل جامعة ما:

| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 52 | 99 | 92 | 86 | 84 | 52 | 99 | 92 | 86 | 84 | 63 | 72 | 76 | 95 | 88 | 63 |
| 72 | 76 | 95 | 88 | 92 | 58 | 65 | 79 | 80 | 92 | 58 | 65 | 79 | 80 | 90 | 75 |
| 74 | 56 | 99 | 90 | 75 | 74 | 56 | 99 | 65 | 79 | 80 | 90 | 75 | 74 | 56 | 99 |
| 90 | 75 | 74 | 56 | 74 | 56 | 99 | 65 | 79 | 80 | 90 | 75 | | | | |

فنجد أنَّ متوسط التكلفة لإصلاح الحاسب $\overline{x}$ يساوي:

$$\overline{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{52+99+92+\ldots\ldots+90+75}{60} = \frac{4681}{60} = 78.02$$

فلو افترضنا أنَّ الأجهزة المعطّلة في الجامعة تمثّل مجتمعاً إحصائياً، فعندئذ نجد أنَّ التباين لتكلفة الإصلاح هو:

$$\sigma^2 = \frac{1}{N^2}\left[N \sum_{i=1}^{N} x_i^2 - \left(\sum_{i=1}^{N} x_i\right)^2\right] = \frac{1}{60^2}\left[60\ (376247) - (4681)^2\right] = 184.18$$

لاحظ هنا أنّنا لا نذكر الوحدة المستخدمة من أجل البيانات لأنّه ليس لها معنى، ففي حالتنا هذا ما معنى أن نقول إنَّ تباين تكلفة إصلاح الأجهزة يساوي 184.18 دولاراً مربّعاً، ولذلك نلجأ إلى حساب قيمة الانحراف المعّياري حيث نجد أنَّ قيمة الانحراف المعّياري لتكلفة الإصلاح تساوي:

$$\sigma = +\sqrt{\sigma^2} = +\sqrt{184.18} = 13.57\ \$$$

وأمّا إذا نظرنا إلى الأجهزة المعطّلة في الجامعة على أنّها عيِّنة من مجتمع الأجهزة المعطّلة، فعندئذ نجد أنَّ التباين لتكلفة الإصلاح هو:

$$S^2 = \frac{1}{n\,(n-1)}\left[n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2\right] = \frac{1}{60 \times 59}\left[60\ (376247) - (4681)^2\right] = 187.305$$

ومن ثمَّ تكون قيمة الانحراف المعّياري لتكلفة الإصلاح تساوي:

$$S = +\sqrt{S^2} = +\sqrt{187.305} = 13.69\ \$$$

٢- بالعودة إلى بيانات المثال / ١/ من (١ , ١ , ٤ , ٢) المُعطاة من خلال الجدول (٩ , ٢) سنقوم بحساب الانحراف المعّياري للأجر السّاعي لعمال كل من المهنتين.

لدينا متوسط الأجر السّاعي للعمال في كل المهنتين يساوي 8 دولارات، وحجم العيِّنة في كل منهما يساوي 8 أشخاص، وبما أنَّ عدد البيانات صغير، فعندئذ:

أ - من أجل البيانات $X$ يُحسب تباينها من خلال العلاقة:

$$S_X^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2 = \frac{1}{7} \sum_{i=1}^{8} \left(x_i - \overline{x}\right)^2 = \frac{90}{7} = 12.86$$

ومن ثمَّ تكون قيمة الانحراف المعّياري للأجر الساعي لعمال المهنة $X$ تساوي:

$$S_X = +\sqrt{S_X^2} = +\sqrt{12.86} = 3.59\ \$$$

ب- وأمّا من أجل البيانات $Y$ فيُحسب تباينها من خلال العلاقة:

$$S_Y^2 = \frac{1}{m-1} \sum_{i=1}^{m} \left(y_i - \overline{y}\right)^2 = \frac{1}{7} \sum_{i=1}^{8} \left(y_i - \overline{y}\right)^2 = \frac{104}{7} = 14.86$$

ومن ثمَّ تكون قيمة الانحراف المعيّاري للأجر السّاعي لعمال المهنة $Y$ تساوي:

$$S_Y = +\sqrt{S_Y^2} = +\sqrt{14.86} = 3.85 \quad \$$$

هكذا نلاحظ أنَّ هذا المقياس أوضح لنا أنَّ تشتّت البيانات $X$ حول متوسطها أقل من تشتّت البيانات $Y$ حول متوسطها. فإذا رجعنا إلى قيم الانحراف المتوسط لهذه البيانات حيث لدينا $\mathbf{MD}_X = 2.50$ و $\mathbf{MD}_Y = 2.25$ فإنَّنا نلاحظ أنَّ مقياس الانحراف المعيّاري أظهر لنا نتيجةً مناقضةً لمقياس الانحراف المتوسط. ولهذا نأخذ بنتيجة الانحراف المعيّاري للبيانات لأنَّها الأكثر دقةً. بالطبع هذه المصادفة في النتائج ليست قاعدة عامّة فمن الممكن أن تتفق نتائج المقياسين مع بعضها.

٣- بالرجوع إلى المثال / ٢/ من (١ , ٢ , ١ , ٢) (قياسات الضغطُ الانقباضي للقلب) حيث لدينا بيانات عيِّنة مقدَّمة من خلال جدول التوزيع التّكراري (٢ , ٢)، فنجد أنَّ التباين لبيانات ذلك الجدول تساوي:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} f_i \left(x_i - \overline{x}\right)^2 = \frac{6602.5}{49} = 134.745$$

ومن ثمَّ تكون قيمة الانحراف المعيّاري للضغط الانقباضي الخاص بالعيّنة التي تمَّ فحص مرضاها تساوي:

$$S = +\sqrt{S^2} = +\sqrt{134.745} = 11.61 \quad \text{mmHg}$$

### (٣ , ٤ , ٢) قاعدة تشبيشيف التجريبية Chebyshev's Empirical Rule

لقد اكتشف عالم الرياضيات الروسي (السوفياتي سابقاً) تشبيشيف Pafnuty Lvovich Chebyshev (1821–1894) أنَّ عدد القياسات التي تقع بين قيمتين متناظرتين حول المتوسط يرتبط بقيمة الانحراف المعيّاري لهذه القياسات. إنَّ قاعدة تشبيشيف التجريبية التي سنقدِّمها بعد قليل (والتي يمكن استنتاجها من متباينة تشبيشيف في نظرية الاحتمالات - انظر المبرهنة (١ , ٥ , ٢ , ٧) في الفصل السابع-) تُعطينا تقديراً متحفظاً (لأنَّها لا تعطي التقدير من أجل كل القيم) لعدد القياسات الواقعة ضمن انحراف قَدره $\pm\, k \cdot S$ عن المتوسط $\overline{x}$، وبذلك يُمكننا تقدير عدد القياسات التي يُفترض أن تنتمي إلى الفترة $[\overline{x} - k\,S\,,\,\overline{x} + k\,S]$. نشير إلى أنَّ هذه القاعدة تكون أقرب إلى الدقّة كلما كان شكل توزيع البيانات أقرب إلى التناظر (أو التماثل).

#### (١ , ٣ , ٤ , ٢) نص القاعدة التجريبية لـ تشبيشيف

لتكن لدينا مجموعة بيانات مُعطاة، و $k > 1$ عدداً حقيقياً مُعطى، فعندئذ إذا كانت هذه البيانات:

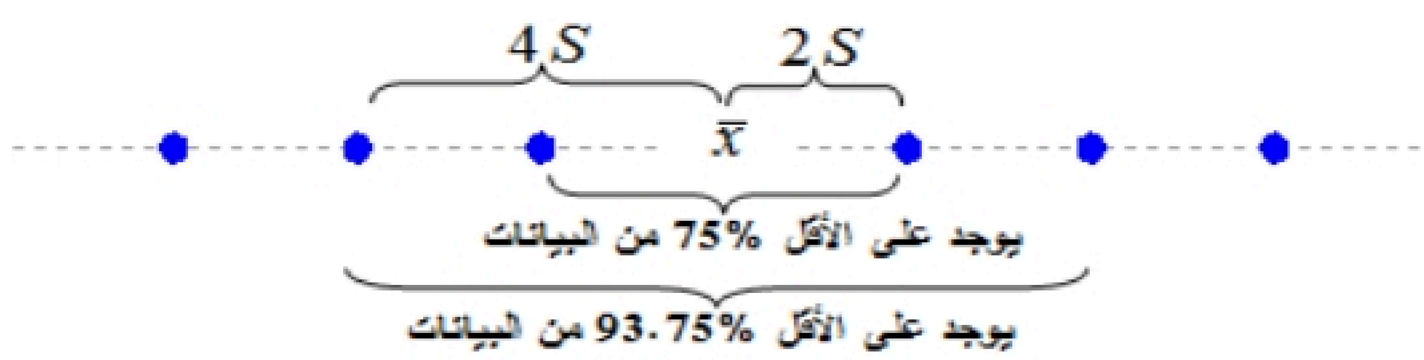
أ - هي $x_1$، $x_2$، ...، و $x_n$ لعيِّنة متوسطها $\overline{x}$ وانحرافها المعيّاري $S$، فإنَّه يوجد على الأقل $\left(1-\frac{1}{k^2}\right) \times 100\%$ من البيانات ضمن الفترة $[\overline{x} - k\,S\,,\,\overline{x} + k\,S]$، ومن ثم يقع $\frac{100}{k^2}\%$ من البيانات خارج هذه الفترة.

ب- هي $x_1$، $x_2$، ...، و $x_N$ لمجتمع متوسطه $\mu$ وانحرافه المعيّاري $\sigma$، فإنَّه يوجد على الأقل $\left(1-\frac{1}{k^2}\right) \times 100\%$ من البيانات ضمن الفترة $[\mu - k\,\sigma\,,\,\mu + k\,\sigma]$، ومن ثم يقع $\frac{100}{k^2}\%$ من البيانات خارج هذه الفترة.

الجدول الآتي والرسم الذي يليه يوضّحان لنا ما سبق من أجل بعض القيم المختلفة لـ $k>1$ (حيث يُلاحظ عدم جدوى هذه القاعدة عندما يكون $k=1$).

الجدول (٢,١١)

| $k$ | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|
| $\frac{1}{k^2}\times 100$ | 25 | 11.11 | 6.25 | 4 | 2.78 | 2.04 | 1.56 | 1.23 | 1 |
| $\left(1-\frac{1}{k^2}\right)\times 100$ | 75 | 88.88 | 93.75 | 96 | 97.96 | 97.96 | 98.44 | 98.77 | 99 |




الشكل (٢,٧)

### (٢,٣,٤,٢) مثال

بالرجوع إلى بيانات المثال / ١/ من (٢,١,٢,١) (مجتمع أطوال طلاب في مدرسة ثانوية). حيث وجدنا متوسط القيم للبيانات المُعطاة $\mu=172.82$، وبحساب قيمة الانحراف المعيّاري لتلك البيانات نجدها (على افتراض أنَّها بيانات مجتمع إحصائي) $\sigma=10.32$، ومن ثَمَّ يُفترض أن يكون لدينا:

١- لا يقل عن 75% من البيانات تقع ضمن الفترة $[\mu-2\sigma\ ,\ \mu+2\sigma]$ التي هي:

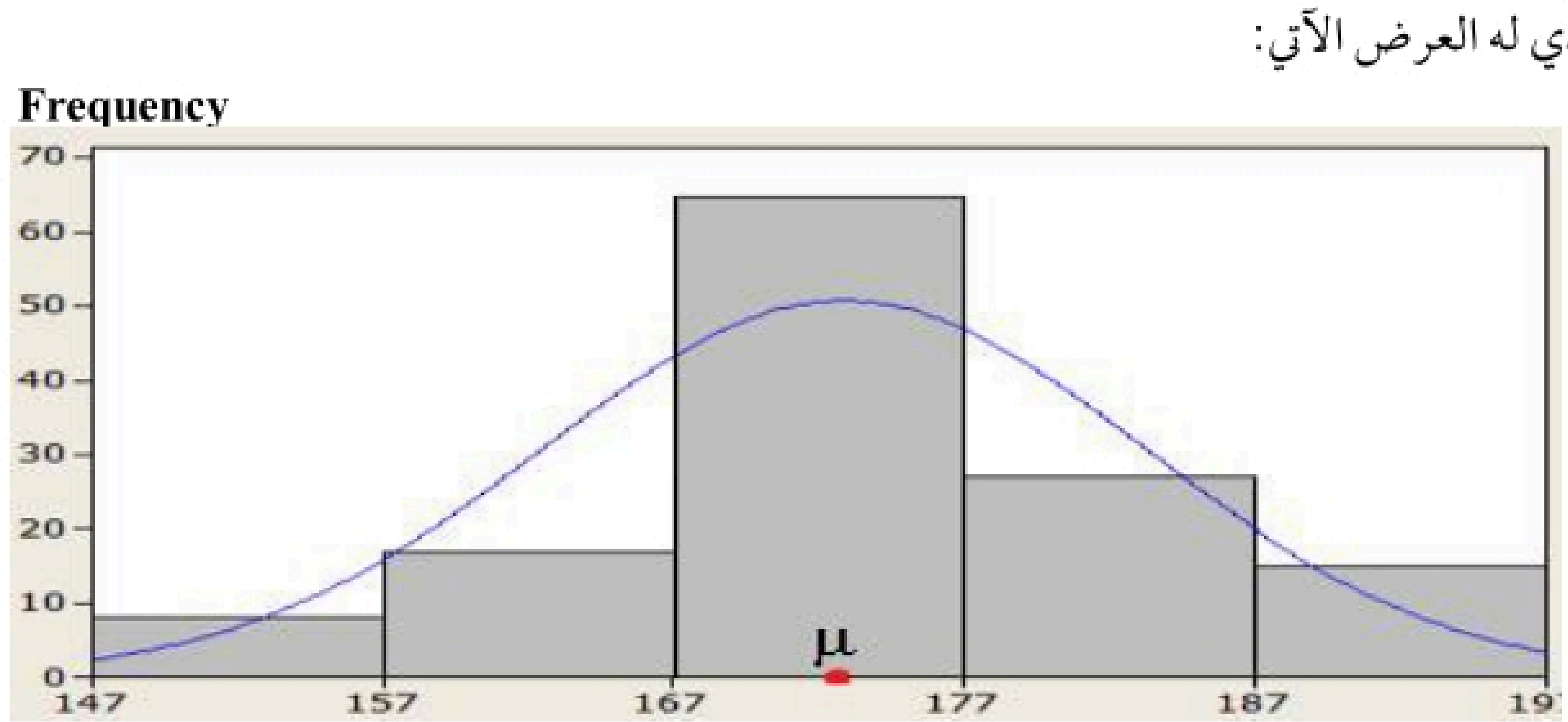
$$[152.18\ ,\ 193.46]=[\ 172.82-20.64\ ,\ \ 172.82+20.64]$$

فلو قمنا بعدّ البيانات الواقعة ضمن هذه الفترة لوجدنا عددها يساوي 123 قيمة، ونسبتها إلى الكل (وهو يساوي 132) يساوي 93.18% وهذا متوافق تماماً مع هذه القاعدة.

٢- لا يقل عن 93.75% من البيانات تقع ضمن الفترة: $[\mu-4\sigma\ ,\ \mu+4\sigma]$ التي هي:

$$[172.82-41.28\ ,\ 172.82+41.28]=[131.54\ \ ,\ \ 214.10\ ]$$

وبعدِّ البيانات الواقعة ضمن هذه الفترة فإنّنا نجدها تساوي 132 قيمة (جميع البيانات) ونسبتها إلى الكل يساوي 100% وهذا متوافق تماماً مع هذه القاعدة أيضاً، وسبب هذا التوافق الممتاز في النتائج يعود إلى توزيع البيانات التي لدينا، فلو قمنا بعرض المدرج التّكراري لتلك البيانات الذي له العرض الآتي:

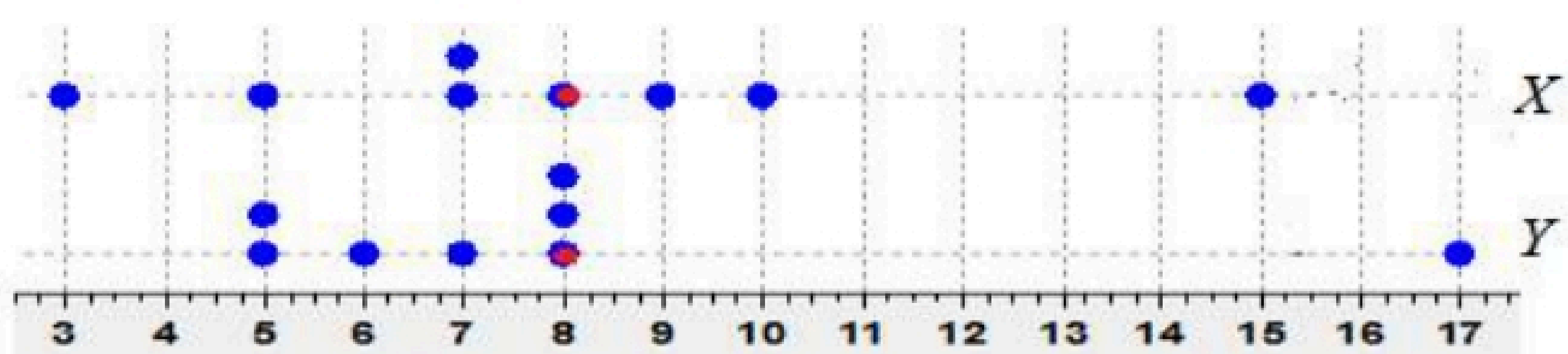


الشكل (٢ , ٨)

فنلاحظ أنَّ لهذا الشكل تناظراً مقبولاً (قريب من التوزيع الطبيعي).

### (٢ , ٤ , ٤) المدى Range

الآن بالعودة إلى مقاييس التشتّت، فإذا صادفنا في بعض الحالات فقداناً لبعض البيانات، فعندئذ لا يمكننا استخدام مقياسي التشتّت السابقين (الانحراف المتوسط والانحراف المعيّاري) لأنّهما يعتمدان في حسابها على جميع قيم البيانات، ومن هنا جاءت أهمية البحث في تقديم مقاييس أخرى للتشتّت لا تعتمد على جميع قيم البيانات. من المقاييس التي يمكن أن تفيدنا في هذا المضمار المدى (الذي قدمنا تعريفه سابقاً لدى تشكيل جدول التوزيع التكراري) وذلك عندما تكون أكبر وأصغر قيمة في البيانات ليست في عداد البيانات المفقودة. إنَّ هذا المفهوم (المدى) هو من المقاييس المستخدمة للدلالة على تشتّت البيانات أيضاً، ولكن لا يُنظر إليه كمقياس جيد للتشتّت رغم أنَّه يعطي صورة عن مدى تشتّت مجموعة من البيانات، إذ إنَّه وفي كثير من الحالات (وخاصة لدى العينات الكبيرة الحجم) لا يُظهر لنا بجلاء توضّع البيانات حول متوسطها لأنَّه يأخذ الفرق بين أكبر وأصغر قيمة فقط، وحتى من أجل بعض العينات الصغيرة فقد لا يكون منصفاً في قياس التشتّت بين البيانات حول متوسطها، فعلى سبيل المثال لو عدنا إلى المثال/ ١/ من (٢ , ٤ , ١ , ١) حيث لدينا مجموعتي بيانات مُعطاة بالجدول (٢ , ٩)، فلو قمنا بحساب المدى للأجر الساعي لعمال كل من المهنتين لوجدنا $\mathcal{R}_X = 15-3=12$ وكذلك $\mathcal{R}_Y = 17-5=12$ أيضاً (انظر الشكل الآتي لانتشار البيانات حول المتوسط لكل من مجموعتي البيانات $X$ و $Y$).




الشكل (٢ , ٩)

علماً أنَّ قيم المتوسط لمجموعة البيانات $X$ و$Y$ هي على الترتيب $\overline{x} = 8$ و $\overline{y} = 8$ وهما متساويتان أيضاً، ولكن نلاحظ تفاوتاً واضحاً في تشتّت البيانات حول متوسطها لكلٍّ من هاتين المجموعتين.

#### (٢ , ٤ , ٤ , ١) مزايا وعيوب المدى كمقياس للتشتّت:

بالرغم من أنَّ ا لمدى يُعدَّ مقياساً ضعيفاً للتشتّت إلا أنَّه يتمتع بمجموعة من المزايا تجعل منه مقياساً مرغوباً في بعض الحالات، فمن مزايا المدى:

١ – إنه بسيط وسهل الحساب.

٢ – يكثر استخدامه عند الإعلان عن حالات الطقس، والمناخ الجوي، مثل درجات الحرارة، والرطوبة، والضغط الجوي.

٣ – له استخدامات مهمّة في بعض الدراسات الإحصائية مثل ضبط الجودة الإحصائي.

#### (٢ , ٤ , ٤ , ٢) عيوب المدى كمقياس للتشتّت:

١ – يعتمد على قيمتين فقط، ولا يأخذ جميع القيم في الحسبان، ومن ثَّم تكون قيمته أقل دقةً من المقاييس السّابقة.

٢ – يتأثر بالقيم المتطرفة بشكل كبير جداً لأنَّه يعتمد على القيم التي تقع على الأطراف أصلاً.

٣ – يصبح عديم الاستخدام إذا فُقِدت أصغر أو أكبر قيمة في البيانات (هنا "أو" لا تفيد الحصر).

إذاً والحال كذلك، فلا بدَّ من البحث عن مقياس آخر يتجاوز السلبيات للمقياس السّابق. إنَّ مقياس التشتّت الآتي يقدّم لنا حلاً لمشكلة وجود قيم متطرّفة في البيانات.

**(٥ , ٤ , ٢) تعريف ( المدى الرُّبَيْعيّ Interquartile Range)**

لتكن لدينا بيانات مُعطاة، فعندئذ يُعرَّف المدى الرُّبَيْعيّ (ويُرمز له IQR ) من خلال العلاقة الآتية:

$$\mathbf{IQR} = \mathbf{Q}_3 - \mathbf{Q}_1 \qquad [2,33]$$

علماً أنَّ $\mathbf{Q}_1$ و $\mathbf{Q}_3$ هما الرُّبَيْعيّ الأول والثالث للبيانات على الترتيب.

لاحظ أنَّ هذا المقياس يَحسب المدى لنصف عدد البيانات الواقعة في الوسط بعد ترتيبها تصاعديًّا، ومن ثمَّ فإنَّ القيم المتطرفة ستصبح خارج نطاق هذا المقياس. من جهة أخرى يُعدّ هذا المقياس (المدى الرُّبَيْعيّ) أفضل من سابقه (المدى) إلاّ أنَّه يعتمد في قراره على قيمتين من البيانات فقط ولا يأخذ في الحسبان مواضع جميع البيانات، ولهذا السبب يُنظر إلى هذا المقياس على أنّه مقياس ضعيف للتشتّت أيضاً، ولكنَّه يُعدّ مقبولاً في حال فقدان بعض البيانات المعلوم ترتيبها وغير الموافقة لقيمتي الرُبيعيين الأول والثالث.

فلو عدنا إلى المثال / ١/ من (١ , ١ , ٤ , ٢)، فإنّنا نجد ما يلي:

$$\mathbf{Q}_{1,X} = 5.50 \;,\; \mathbf{Q}_{1,Y} = 5.25 \;,\; \mathbf{Q}_{3,X} = 9.75 \text{ and } \mathbf{Q}_{3,Y} = 8$$

وهكذا يكون لدينا المدى الرُّبَيْعيّ لمجموعة البيانات $X$ هو:

$$\mathbf{IQR}_X = \mathbf{Q}_{3,X} - \mathbf{Q}_{1,X} = 9.75 - 5.50 = 4.25$$

في حين نجد أنَّ المدى الرُّبَيْعيّ لمجموعة البيانات $Y$ هو:

$$\mathbf{IQR}_Y = \mathbf{Q}_{3,Y} - \mathbf{Q}_{1,Y} = 8 - 5.25 = 2.75$$

أي إنَّ لمجموعة البيانات $X$ تشتّت مختلف عن تشتّت مجموعة البيانات $Y$ ، وهكذا نجد أنّه ما عجز عنه المدى (في إظهار فروق التشتّت لمجموعتي البيانات ) أنجزه المدى الرُّبَيْعيّ، حيث بيَّن وجود فارق بين تشتّت البيانات لكلٍّ من هاتين المجموعتين.

الآن، وقبل الانتقال للفقرة التالية نشير إلى أنَّ نصف قيمة المدى الرُّبَيْعيّ(وسنرمز لها بـ Q ) تُدعى الانحراف الرُّبَيْعيّ Quartile Deviation، وتُستخدم كمقياس للتشتّت أيضاً. أي إنَّ:

$$\mathbf{Q} := \frac{\mathbf{Q}_3 - \mathbf{Q}_1}{2} \qquad [2,34]$$

لقد لاحظنا أنَّ مقاييس التشتّت تعتمد على وحدة القياس المستخدمة من أجل البيانات، ومن ثَّم يكون من الصعب علينا (إن لم يكن متعذّراً) إجراء المقارنة بين التشتّت لمجموعتين أو أكثر من البيانات عندما تكون وحدات القياس المستخدمة من أجل كلٍّ منها مختلفة عن الأخرى، لذلك كان لابدَّ من وضع معيّار يمكننا من إجراء المقارنة بين تشتّت البيانات حول متوسطها لمجموعتين أو أكثر من البيانات حتى في حال كانت وحدات القياس مختلفة بعضها عن بعض. إنَّ المعيّار الآتي والمُسمَّى *مُعامل التغيُّر* يُحقِّق لنا ذلك، إذ إنَّه يستخدم نسبة الانحراف المعيّاري إلى المتوسط الحسابي لإلغاء خاصية وحدة القياس المستعملة وإظهار مفعول التباين بآن واحد، ومن ثَّم تقديم هذه القيمة كنسبة مئوية خاصة بالبيانات.

**(٦ , ٤ , ٢) تعريف (مُعامل التَغيُّر Coefficient of Variation)**

إنّ مُعامل التغيُّر يقدِّم لنا قيمة تجعلنا نشعر بمدى التغيّر الحاصل لمتغيِّر ما، وهو يُمثِّل نسبة الانحراف المعيّاري إلى المتوسط، ويفيدنا في مقارنة إحصائيّة لدرجة التباين من مجموعة بيانات إلى أخرى حتى ولو كانت وحدات القياس تختلف بشكلٍ كبيرٍ بعضها عن بعض، وذلك لأنَّ قيمة هذا المُعامل هي عدد مجرَّد، ومن أجل تعريفه سنميِّز بين حالتين:

أ- إذا كانت البيانات المُعطاة لعيِّنة بمتوسط $\overline{x} \neq 0$ وانحراف معياري $S$ ، فعنـدئذ يُعرَّف **مُعامل التَغيُّر** (الذي يُرمز له بـ**CV**) لهذه البيانات من خلال العلاقة الآتية:

$$\mathbf{CV} := \frac{S}{\overline{x}} \times 100 \ \% \qquad [2,35\text{-a}]$$

ب- إذا كانت البيانات المُعطاة لمجتمع بمتوسط $\mu \neq 0$ وانحراف معياري $\sigma$ ، فعنـدئذ يُعرَّف **مُعامل التَغيُّر** (الذي يُرمز له بـ **CV** أيضاً) لهذه البيانات من خلال العلاقة الآتية:

$$\mathbf{CV} := \frac{\sigma}{\mu} \times 100 \ \% \qquad [2,35\text{-b}]$$

وفي كلا الحالتين يُقرأ الناتج كنسبة مئوية (كما هو واضع في العلاقتين السّابقتين).

### (٢ , ٤ , ٦ , ١) مثال

لتكن لدينا مجموعتي بيانات تمثِّل الطول والوزن لستة أشخاص (ذكور) بالغين، ومُقدَّمتين من خلال الجدول الآتي:

الجدول (٢ , ١٢)

| الشخص | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| $X$ الطول بـ سم | 182 | 173 | 168 | 177 | 169 | 189 |
| $Y$ الوزن بـ كغ | 85 | 72 | 68 | 73 | 71 | 93 |

حيث نلاحظ اختلاف وحدة القياس بين هاتين المجموعتين.

الآن، وبحساب قيم المتوسط والانحراف المعيّاري لكل من هاتين المجموعتين نجد أنَّ:

$$\overline{x} = 176.33 \ cm \quad \& \quad S_X = 8.09 \ cm \quad \& \quad \overline{y} = 77 \ Kg \quad \& \quad S_Y = 9.78 \ Kg$$

ومنه تكون قيمة مُعامل التَغيُّر لمجموعة البيانات $X$ هي $\mathbf{CV}_X = 4.59 \ \%$ ، في حين نجد قيمة مُعامل التغيُّر لمجموعة البيانات $Y$ هي $\mathbf{CV}_Y = 12.70 \ \%$ ، ومن ثَمَّ تكون قيمة مُعامل تغيُّر الطول أصغر بكثير من مُعامل تغيُّر الوزن لهذه العيِّنة من الأشخاص.

### (٢ , ٤ , ٦ , ٢) ملاحظات

١- تُعْرَف القيمة المطلقة لمُعامل الاختلاف باسم **الانحراف المعيّاري النسبي** Relative Standard Deviation (**RSD**)، ويُعبَّر عنها كنسبة مئوية أيضاً.

٢- في حال تَعَذُّر حساب المتوسط أو التباين فإنَّه يمكن أن يُستعاض عن مُعامل التغيُّر بمُعامل آخر يُدعى مُعامل التشتّت، وهو يُحسب بدلالة الرُبعيين الأول والثالث، ويقدِّمه لنا التعريف الآتي.

### (٢ , ٤ , ٧) تعريف (مُعامل التشتّت Coefficient of Dispersion)

لتكن لدينا مجموعة بيانات مُعطاة لها رُبيعي أول $\mathbf{Q}_1$ ورُبيعي ثالث $\mathbf{Q}_3$، فعندئذ يُعرَّف **مُعامل التشتّت** (الذي يُرمز له بـ **CD**) من خلال العلاقة الآتية:

$$\mathbf{CD} := \frac{\mathbf{Q}_3 - \mathbf{Q}_1}{\mathbf{Q}_3 + \mathbf{Q}_1} \times 100\ \% \qquad [2,36]$$

**(٢, ٤, ٧, ١) مثال**

لتكن لدينا مجموعة بيانات خام مُرتّبة، ولكنَّها تحوي بيانات مفقودة معلوم ترتيبها ومُعطاة كما في العرض الآتي:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 501 | ? | 497 | 495 | 493 | 489 | 485 | ? | 482 | ? |
| 560 | 541 | ? | 537 | 528 | 520 | 515 | 513 | 511 | 505 |

فنجد أنَّ حساب مُعامل التغيُّر لهذه البيانات غير مُمكن بسبب عدم إمكانية حساب كلٍّ من المتوسط ومن ثمَّ الانحراف المعّياري (لاحظ هنا حتى المدى غير ممكن الحساب)، وأمَّا لحساب مُعامل التشتّت، فنجد أنَّ قيمة الرُبّيعيّ الأول والثالث تساوي:

$$\mathbf{Q}_1 = x_5 + \tfrac{1}{4}(x_6 - x_5) = 489 + \tfrac{1}{4}(493 - 489) = 450$$

$$\mathbf{Q}_3 = x_{15} + \tfrac{3}{4}(x_{16} - x_{15}) = 520 + \tfrac{3}{4}(528 - 520) = 526$$

ومن ثمَّ (بحسب العلاقة [2,36]) قيمة مُعامل التشتّت لهذه البيانات تساوي:

$$\mathbf{CD} = \frac{526 - 450}{526 + 450} \times 100 = 7.79\,\%$$

الآن، وقبل ختام بحث مقاييس التشتّت سنقدِّم مفهوماً متعلِّقاً بقيمة المتوسط والانحراف المعّياري لمجموعة بيانات، ويُدعى **الدرجة المعّيارية** Standard Score حيث ينطوي تحت هذه التسمية معيارين مهمّين هما $Z$-score و$T$-score، وسنوضّحهما من خلال الفقرات الآتية.

**(٢, ٤, ٨) تعريف (الدرجة المعّيارية $Z$ Z-score)**

من أجل تعريف الدرجة المعّيارية $Z$ سنميِّز بين حالتين:

أ- إذا كانت $x_1$، $x_2$، ...و $x_n$ بيانات مُعطاة لعيِّنة بمتوسط $\overline{x}$ وانحراف معّياري $S$، فعندئذ **الدرجة المعّيارية** $Z$ للقيمة $x_i$ من هذه البيانات (والتي سنرمز لها بـ $z_i$) تُعرَّف من خلال العلاقة الآتية:

$$Z_i := \frac{x_i - \overline{x}}{S} \quad ; i \in \mathbb{N}_n \qquad [2,37\text{-}a]$$

ب- إذا كانت $x_1$، $x_2$، ...و $x_N$ بيانات مُعطاة لمجتمع بمتوسط $\mu$ وانحراف معياري $\sigma$، فعندئذ **الدرجة المعّيارية** $Z$ للقيمة $x_i$ من هذه البيانات تُعرَّف من خلال العلاقة الآتية:

$$Z_i := \frac{x_i - \mu}{\sigma} \quad ; i \in \mathbb{N}_n \qquad [2,37\text{-}b]$$

لاحظ أنَّ قيمة هذا المعّيار هي مُؤشِّر على انحراف قيمة البيان الخام عن المتوسط، ومن ثمَّ فإنَّها تُحدِّد موقع البيان الخام من المتوسط اتّجاهاً وبُعداً، فالاتّجاه تُحدِّده إشارة (- أو+)، فإذا كانت قيمته موجبة فإنَّ ذلك يعني أنَّه أكبر من المتوسط والعكس إذا كانت قيمته سالبة، وأمّالطُبُفتعني كـ بَر القيمة المطلقة له، فكلما كَبُرت قيمته المطلقة ابتعدت القيمة $x_i$ عن المتوسط.

من فوائد الدرجة المعّيارية أنّها تُعطينا صورة عن مكان الدرجة من المتوسط كما ذكرنا ذلك آنفًا، فعلى سبيل المثال يمكننا التعرّف على مستوى أداء طالب بالنسبة إلى زملائه، ولتوضيح ذلك سنقدِّم المثال الآتي.

**(١ , ٨ , ٤ , ٢) مثال**

في إطار الاختبارات العامة قامت وزارة التربية باختبار تدريبي لطلاب الصف الثالث الثانوي، ومن ثمَّ أُخذت عيِّنة من مدرستين $X$ و $Y$ . بعد ذلك سُحِبَت من كل مدرسة نتيجة لطالب واحد، فوجد أنَّ الطالب $A$ من المدرسة $X$ قد حصلَ على 88 درجة، في حين أنَّ الطالب $B$ من المدرسَة $Y$ قد حصل على 83 درجة، فهل يُمكننا الادِّعاء أنَّ أداء الطالب $A$ أفضل من أداء الطالب $B$ ؟

في الواقع إنَّ الإجابة على هذا السؤال يتطلب معرفة مستوى كل من هذين الطالبين في مدرسته لأنَّهما لا يخضعان لتعليم موحَّد سواء من حيث البيئة المحيطة بالطالب (من تفاعل ونشاط، أو من حيث المعلِّم، وحتى لو كانا من مدرسة واحدة فربما لم يُدرسا من قبل معلِّم واحد أيضاً)، ولذلك فإنَّ الإجابة على السؤال المطروح ليست بهذه البساطة، ويجب أن يُؤخذ في الحُسبان درجات الاختبار لزملاء كل واحد منهم، ومن ثمَّ يُبحث في الإجابة بناءً على هذه المُعطيات. فلو افترضنا الآن أنَّ درجات الاختبار للطلاب في المدرسة $X$ لها متوسطٌ يساوي 85 درجة بانحراف معياري يساوي 3 درجات، وأنَّ درجات الاختبار للطلاب في المدرسة $Y$ لها متوسط يساوي 72 درجة بانحراف معياري يساوي 5 درجات، فعندئذ بحساب الدرجة المعيّارية $z$ لكل من هذين الطالبين نجد الآتي:

$$z_A = \frac{88-85}{3} = 1 \qquad \& \qquad z_B = \frac{81-72}{5} = 1.8$$

وهكذا نجد أنَّ الدرجة المعيّارية $z$ للطالب $B$ أعلى من الدرجة المعيّارية $z$ للطالب $A$ بفارق كبير (تكاد تقارب الضعف)، وهذا يعني أنَّ أداء الطالب $B$ أفضل من أداء الطالب $A$ رغم أنَّ نتيجة الطالب $B$ أقل من نتيجة الطالب $A$. لهذا السبب قد تقوم وزارات التعليم العالي في بعض الدول بقبول معدَّلات (أو مجموع درجات) أقل من المطلوب للطلاب الناجحين في الثانوية العامّة من المناطق النائية من أجل التسجيل في اختصاص ما، وذلك للأسباب التي أوضحنا بعضها في المثال السابق. من هنا يتبيَّن لنا الفارق بين النتيجة الامتحانية والنتيجة المعيّارية كمقياس للتفوق، وكذلك يتضَّح لنا من هذا المثال أنَّ الدرجة المعيّارية $z$ تفيدنا بالموضع النسبي لقيمة من مجموعة قيم مفترضة، وبما أنَّها قيمة مجرَّدة من وحدة القياس فإنَّها تصلح للمقارنة بين الدرجات المعيّارية $z$ لمجموعتين أو أكثر من البيانات التي لكلٍّ منها وحدة قياس مختلفة عن الأخرى.

**(٢ , ٨ , ٤ , ٢) ملاحظات**

١- لو حُسبت الدرجة المعيّارية $z$ لجميع عناصر العيِّنة (أو المجتمع) فعندئذ سنلاحظ أنَّ البيانات النّاتجة عن هذه الدرجات المعيّارية $z$ لها متوسط يساوي الصفر وانحراف معياري يساوي الواحد. في الحقيقية إذا أُجريت عمليّات حسابيّة على قيمة عددية $x$ بحيث يُصبح للقيمة الجديدة $z$ (بعد تنفيذ العمليات الحسابية على $x$) متوسط يساوي الصفر وانحراف معياري يساوي الواحد، فعندئذ يُقال عن هذه العمليّات التي تمّ تطبيقها على $x$ إنَّها **عمليّة استعيار** Standardization للقيمة العددية $x$، بمعنى أنَّ تحويل قيمة عددية $x$ إلى قيمة $z$ لها متوسط يساوي الصفر وانحراف معياري يساوي الواحد تُدعى عمليّة استعيّار، وهذه التسميّة تبقى سارية المفعول حتى من أجل المتغيِّرات العشوائية وغير العشوائيّة أيضاً (انظر استعيّار المتغيّر العشوائي الطبيعي في الفصل السادس).

٢- يُأخذ على الدرجة المعيّارية $z$ أنَّها عند قيمة الصفر يكون للبيان الذي استعيِّرت قيمته درجة القيمة الوسطى $\bar{x}$ (أو $\mu$)، ولذلك قد لا يكون معناها واضحاً لدى الكثير، وعلاوة على ذلك فقد تكون هذه الدرجة المعيّارية $z$ سالبة أيضاً، وهذا بدوره يجعل تفسيرها من أجل بعض الحالات غير واضح أو غير ذي معنى، ولذلك اقترح تقديم مقياس آخر يتجاوز هذه السلبية.

إنَّ الدرجة المعّيارية البديلة للدرجة المعّيارية $\mathcal{Z}$ تُعرف باسم "**الدرجة المعّيارية** $\mathcal{T}$ " $\mathcal{T}$*-score* وتُعرف باسم **الدرجة التّائيّة** (أخذ الحرف $\mathcal{T}$ من اسم واضع هذاالمعيّار عالم النفس الأمريكي ثورنيدك **(Edward Lee Thorndike (1874-1949)** أيضاً والتي يُصيغها لنا التعريف الآتي.

**(٢, ٤, ٩) تعريف (الدرجة المعّيارية $\mathcal{T}$ $\mathcal{T}$-score)**

من أجل تعريف الدرجة المعّيارية $\mathcal{T}$ سنميِّز بين حالتين:

أ - إذا كانت $x_1$، $x_2$،...،و $x_n$ بيانات مُعطاة لعيِّنة بمتوسط $\overline{x}$ وانحراف معياري $S$، فعندئذ **الدرجة المعّيارية** $\mathcal{T}$ لقيمة $x_i$ من هذه البيانات (التي سنرمز لها بـ $\mathcal{T}_i$ ) تُعرَّف من خلال العلاقة الآتية:

$$\mathcal{T}_i := 10\left(\frac{x_i - \overline{x}}{S}\right) + 50 \quad ; i \in \mathbb{N}_n \qquad [2,38\text{-a}]$$

ب- إذا كانت $x_1$، $x_2$،...،و $x_N$ بيانات مُعطاة لمجتمع بمتوسط $\mu$ وانحراف معياري $\sigma$ ، فعندئذ تُعرَّف **الدرجة المعّيارية** $\mathcal{T}_i$ للقيمة $x_i$ من خلال العلاقة الآتية:

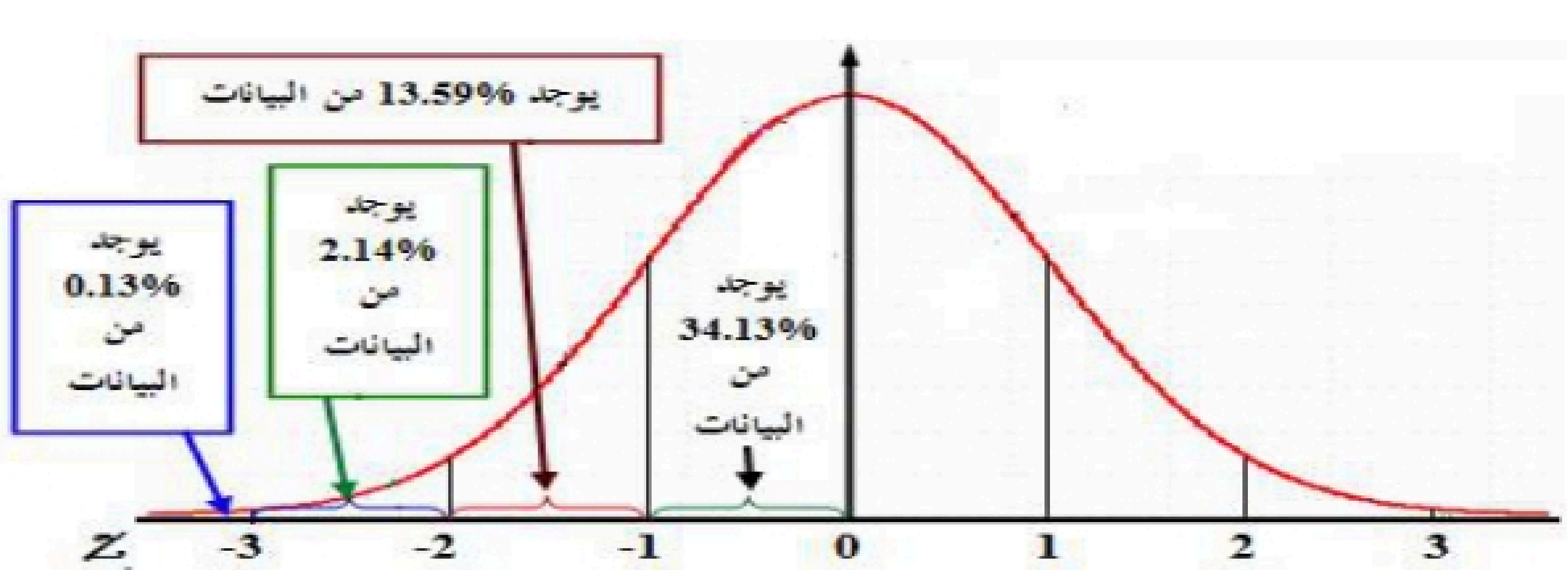
$$\mathcal{T}_i := 10\left(\frac{x_i - \mu}{\sigma}\right) + 50 \quad ; i \in \mathbb{N}_n \qquad [2,38\text{-b}]$$

**(٢, ٤, ٩, ١) ملاحظات**

١- من التعريف السابق للقيمة التّائيّة يتضح لنا أنَّه عندما تكون قيمة $\mathcal{Z} = 0$ فإنَّ $\mathcal{T} = 50$ ، ومن ثَّم نكون قد تخلصنا من قيمة الصفر كقيمة لهذا المعّيار علاوةً على تجنب الكسور (أو التقليل منها) بسبب الضرب بالعدد 10، والشكل الآتي (٢, ١٠) يوضِّح لنا ذلك، ويظهر قيم كل من $\mathcal{Z}$ و $\mathcal{T}$ وما يقابلها من نسب عدد البيانات وفقاً لقاعدة تشبيشيف التجريبيّة على منحنى متناظر ومعتدل.

٢- يُلاحظ أنَّ لـ $\mathcal{T}$ (كمتغيِّر أو كدالَّة بالبيانات) متوسط يساوي 50 وانحراف معياري يساوي 10، وذلك لأنَّ لـ $\mathcal{Z}$ (كمتغيِّر أو كدالَّة بالبيانات) متوسط يساوي 0 وانحراف معياري يساوي 1.

٣-من الصيغة المقدَّمة للقيمة التّائيّة نجد أنَّ $\mathcal{T} = 10\mathcal{Z} + 50$ وهي دالَّة خطية بـ $\mathcal{Z}$ ، ومن ثَّم $\mathcal{T}$ هو تحويل خطي.




الشكل (٢, ١٠)

لتوضيح دلالة هذا المفهوم سنقدِّم المثال الآتي.

**(٢, ٤, ٩, ٢) مثال**

بالعودة إلى المثال السّابق (١ , ٨ , ٤ , ٢) سنقوم بحساب الدرجة التّائيّة لكلٍّ من الطالبين $A$ و $B$. حيث قمنا سابقاً بحساب كلٍّ من $z_A$ و $z_B$ ووجدنا أنَّ $z_A = 1$ و $z_B = 1.8$ ، ومن ثَمَّ يكون لدينا:

$$\mathcal{T}_A = 10(1) + 50 = 60 \qquad \& \qquad \mathcal{T}_B = 10(1.8) + 50 = 68$$

وهكذا نجد أنَّ الدرجة المعيّارية $\mathcal{T}$ للطالب $B$ أعلى من الدرجة المعيّارية $\mathcal{T}$ للطالب $A$ بفارق واضح (ولكن لم يَعدْ يمكننا الادعاء أنّه يقارب الضعف كما سبق في المعيّار السّابق)، وهذا يعني أنَّ أداء الطالب $B$ أفضل من أداء الطالب $A$ رغم أنَّ نتيجة الطالب $B$ أقل من نتيجة الطالب $A$.

## (٢,٥) مقاييس الشكل للتوزيعات التكرارية

## Form Measures of Frequency Distributions

لقد قدَّمنا في الفصل الأول مناقشة بسيطة وموجزة حول أشكال التوزيعات، ولكن دون التطرق للمقاييس التي تُحدِّد لنا نوعيّة هذه الأشكال. إنَّ المقاييس الجيدة والدقيقة لأشكال التوزيعات تعتمد وبشكل أساسي على مفهوم العزوم المركزيَّة الذي سنقدِّمه على النحو الآتي.

### (١ , ٥ , ٢) العزوم المركزيَّة لمجموعة بيانات Central Moments of Data Set:

من المفاهيم اللازمة لدراسة مقاييس الشكل للتوزيعات التكراريّة ما يُعرف باسم **العزم المركزي من المرتبة $r$** لمجموعة بيانات، وهذا المفهوم يقدَّمه لنا التعريف الآتي.

#### (١ , ١ , ٥ , ٢) تعريف (العزم المركزي من المرتبة $r$ لمجموعة بيانات) Central Moment of Order $r$ of Data Set

ليكن $r$ عدداً صحيحاً موجباً كيفياً ولكن مثَّبتاً، فعندئذ لتعريف العزم المركزي من المرتبة $r$ لبيانات مُعطاة سنناقش الحالتين الآتيتين:

**أولاً: من أجل البيانات الخام:**

أ- إذا كانت $x_1$ ، $x_2$ ، ... و $x_n$ بيانات عيِّنة متوسطها $\overline{x}$ ، فإنَّ **العزم المركزي من المرتبة $r$** لهذه البيانات سنرمز له بـ $\mathbf{S}^{(r)}$ ويُحسب بالعلاقة الآتية:

$$\mathbf{S}^{(r)} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^r \qquad [2,39\text{-a}]$$

ب- إذا كانت $x_1$ ، $x_2$ ، ... و $x_N$ بيانات مجتمع متوسطه $\mu$ ، فإنَّ العزم المركزي من المرتبة $r$ لهذه البيانات سنرمز له بـ $\sigma^{(r)}$ ويُحسب بالعلاقة الآتية:

$$\sigma^{(r)} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^r \qquad [2,39\text{-b}]$$

**ثانياً: من أجل البيانات المُجمَّعة:**

أ - إذا كانت مجموعة البيانات المُعطاة هي بيانات عيِّنة مُجمَّعة (أو مُحمَّلة، أو مبوَّبة، أو مصبوبة) في جدول توزيع تكراري بـ $k$ فئة كما في الجدول (١ , ٢)، وكان متوسطها يساوي $\overline{x}$، فعندئذ يُحسب العزم المركزي $\mathbf{S}^{(r)}$ لهذه البيانات بالعلاقة الآتية:

$$\mathbf{S}^{(r)} = \frac{1}{(\sum f_i) - 1}\sum_{i=1}^{k} f_i \left(x_i - \overline{x}\right)^r \qquad [2,40\text{-a}]$$

ب- إذا كانت مجموعة البيانات المُعطاة هي بيانات مجتمع ومُحمَّلة في جدول توزيع تكراري بـ $k$ فئة كما في الجدول (١ , ٢)، وكان متوسطها $\mu$، فعندئذ يُحسب العزم المركزي $\sigma^{(r)}$ بالعلاقة الآتية:

$$\sigma^{(r)} = \frac{1}{(\sum f_i) - 1} \sum_{i=1}^{k} f_i \ (x_i - \mu)^r \qquad [2,40\text{-}b]$$

### (٢ , ٥ , ١ , ٢) ملاحظات

١- في الحالة الخاصّة عندما يكون $r = 2$ نلاحظ أنَّ العزم المركزي من المرتبة الثانية للبيانات هو التباين لهذه البيانات، ولذلك يُكتب وفي هذه الحالة الخاصّة $\mathbf{S}^{(2)} = S^2$ و $\sigma^{(2)} = \sigma^2$ للدلالة على أنَّ العزم المركزي من المرتبة الثانية هو مقدار غير سالب.

٢- هنا يجب الانتباه إلى أنَّه في الحالة العامّة لدينا $\sigma^{(r)} \neq \sigma^r$ وكذلك $\mathbf{S}^{(r)} \neq S^r$ لأنَّ $0 \le \sigma^r$ وبالمثل $0 \le S^r$ دوماً، في حين أنَّ المقدار $\sigma^{(r)}$ (أو $\mathbf{S}^{(r)}$ ) قد يكون سالباً من أجل القيم الفردية لـ $r$، والمثال الآتي يوضِّح لنا ذلك.

### (٢ , ٥ , ١ , ٣) مثال

لتكن لدينا بيانات عيِّنة مقدَّمةً كما يلي:

9 5 8 4 7 4 8 4 8 4 8 5 6 9 7

فنجد أنَّ قيمة المتوسط لهذه البيانات هو $\overline{x} = 6.4$، ولنقم بحساب $S^5$ و $\mathbf{S}^{(5)}$ مستعينين بالجدول الآتي:

الجدول (٩ , ٢)

| $i$ | $x_i$ | $(x_i - 6.4)^2$ | $(x_i - 6.4)^5$ | $i$ | $x_i$ | $(x_i - 6.4)^2$ | $(x_i - 6.4)^5$ |
|---|---|---|---|---|---|---|---|
| 1 | 9 | 6.76 | 118.8138 | 9 | 8 | 2.56 | 10.48576 |
| 2 | 5 | 1.96 | -5.37824 | 10 | 4 | 5.76 | -79.6262 |
| 3 | 8 | 2.56 | 10.48576 | 11 | 8 | 2.56 | 10.48576 |
| 4 | 4 | 5.76 | -79.6262 | 12 | 5 | 1.96 | -5.37824 |
| 5 | 7 | 0.36 | 0.07776 | 13 | 6 | 0.16 | -0.01024 |
| 6 | 4 | 5.76 | -79.6262 | 14 | 9 | 6.76 | 118.8138 |
| 7 | 8 | 2.56 | 10.48576 | 15 | 7 | 0.36 | 0.07776 |
| 8 | 4 | 5.76 | -79.6262 | sum | 96 | 51.6 | − 49.5456 |

لتسهيل العمليات الحسابية حيث نجد أنَّ قيمة التباين لهذه البيانات تساوي:

$$S^2 = \frac{1}{14} \sum_{i=1}^{15} (x_i - 6.4)^2 = \frac{51.6}{14} = 3.6857$$

ومنه تكون قيمة الانحراف المعيّاري $S = 1.9198$، ومن ثمَّ $S^5 = 26.0798$، وأمّا العزم المركزي من المرتبة 5 لهذه البيانات فنجده يساوي:

$$\mathbf{S}^{(5)} = \frac{1}{14} \sum_{i=1}^{15} (x_i - 6.4)^5 = \frac{-49.5456}{14} = -3.539$$

وهكذا يكون لدينا $\mathbf{S}^{(5)} \neq S^5$، وهذا يؤكد صحة دعوانا.

الآن، وقبل البدء في تقديم بعض مقاييس الشكل يجب أن نتفهَّم السبب وراء استخدام هذه المقاييس، وذلك لأنَّه يجب أن نعلم متى يلجأ الباحث (أو الدارس لمسألة إحصائية ما) إلى حساب مقاييس الشكل.

لاحظنا فيما سبق أنَّ التوزيع يكون ملتوياً إذا تراكمت البيانات عند أحد أطراف التوزيع دون الطرف الآخر، وبالطبع توجد أسباب عديدة لذلك، فعلى سبيل المثال:

- إذا كان اختبار مقرَّر دراسي ما غاية في السهولة أو غاية في الصعوبة فإنَّ قيم الدرجات ستتراكم على الجهة اليُمنى أو اليُسرى على الترتيب، ومن ثمَّ توزيع الدرجات لهذا الاختبار سيكون ملتوياً.
- كذلك بعض المقاييس الفسيولوجية كمقاييس زمن الاستجابة وسرعة الأداء تحتمل أن يكون لها توزيعات ملتوية.
- يمكن لتوزيع البيانات المقاسة على ميزان فتراتي أو رتبي أن يكون ملتوياً أيضاً.

نشير هنا إلى أنَّه إذا كانت البيانات مُجمَّعة في فئات وكان توزيعها ملتوياً فإنَّه يُفضل استخدام الوسيط كمقياس للنَّزعة المركزيَّة، والمدى الرُّبَيْعيّ كمقياس للتشتّت، ولكن بما أنَّ الكثير من الدراسات الإحصائية تفترض أن توزيع الدرجات الخام لمتغيِّر ما يكون معتدلاً ومتناظراً وليس ملتوياً، فإنَّ قيمة مُعامل الالتواء ستكون صفراً أو قريبةً جداً من الصفر، وفي هذه الحالة ينطبق الوسيط على المتوسط ولا داعي لتطبيق مقياس إحصائي ليبين أن التوزيع ليس ملتوياً. لكن الباحث يمكنه تحديد درجة الالتواء ويقرِّر إن كان لابد من إجراء بعض التصحيحات (مثل تصحيح دقة جهاز القياس المستخدم) قبل أن يستمر في تحليل بياناته، وفي حالات أخرى لكي يجعل الباحث توزيع الدرجات قريباً من التناظر (إذا لم يكن كذلك)، فربما يلجأ إلى استخدام نوع من التحويلات غير الخطية على البيانات، غير أن هذه التحويلات تؤدي إلى مشكلات من نوع آخر عند تفسير نتائج البيانات أيضاً.

في الواقع نجد أنَّ طبيعة البحث، ونوع المتغيِّرات موضع الدراسة، وحجم العيِّنة تُعدّ جميعها من العوامل التي يجب أن تأخذ في الحسبان قبل أن يُقرِّر إن كان لابد من حساب مقاييس الالتواء والتفلطح. علاوةً على ما سبق يجب أن يدرك الباحث أنَّ التوزيعات الملتويّة لكثير من المتغيِّرات قد تكون مُصطنعة كما هو الحال لدى الدراسات النفسية والتربوية والاجتماعية، وذلك لأنَّه يندر أن تكون الوحدات المُستخدمة في بناء المقاييس لهذه الدراسات متساوية حيث الكثير من هذه المتغيِّرات تُقاس بعدد من العبارات التي يعطي كل فرد رأيه فيها، أو عدد الإجابات الصحيحة من الأسئلة التي تطرح عليهم. في مثل هذه الحالات يُعيَّن شكل التوزيع الناتج إلى حدٍّ كبير بالنسبة المئوية للعبارات التي أجيب عنها، أو بصعوبة الأسئلة المقدَّمة، فإذا كانت الأسئلة متوسطة الصعوبة فإنَّنا سنتوقع توزيعاً متناظراً لقيم البيانات، وأمَّا إذا كانت الأسئلة سهلةً فإنَّ قيم البيانات ستتراكم عند النهاية العليا للتوزيع، وفي حال كانت الأسئلة صعبةً فإنَّ قيم البيانات ستتراكم عند النهاية السفلى للتوزيع كما ذكرنا ذلك سابقاً.

إذاً، وفي حال غياب وحدات قياس متساوية لأداة القياس لا يمكننا حقيقة القول إنَّ التوزيع متناظر أو ملتوي، وكل ما يمكننا قوله هو أنَّ شكل التوزيع سوف يعتمد على وحدات القياس المستخدمة فقط.

لنعد الآن إلى تقديم بعض العلاقات التي تساعدنا في حساب مقاييس الالتواء والتفلطح لتوزيعٍ ما. حيث نعلم من الفصل السّابق أنَّ أشكال التوزيعات التكرارية تُقسم إلى صنفين:

**الأول**: يتعلَّق بمسألة تناظر شكل التوزيع من عدمه، وهذه تندرج تحت مفهوم الالتواء للتوزيع.

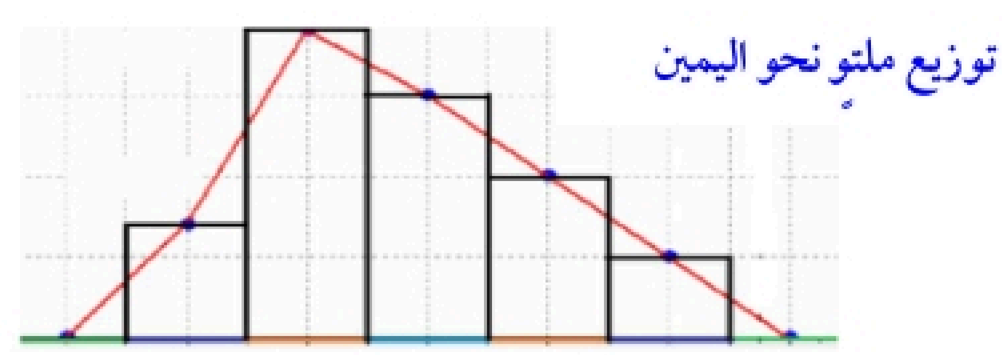
**الثاني**: يتعلَّق بوضع قمة التوزيع، فهل هي منخفضة (أي انبساط شكل التوزيع) أم أنَّها مرتفعة (أي تدبدب شكل التوزيع) أم معتدلة (ليست منخفضة ولا مرتفعة)؟ وهذه الأخيرة تندرج تحت مفهوم التفلطح للتوزيع.

إنَّ المقاييس التي تقدِّم أجوبة على هذه التساؤلات تندرج تحت ما يُعرف باسم **مقاييس الشكل** للتوزيعات التي سنبدأها بمقاييس الالتواء.

### (٢ , ٥ , ٢) مقاييس الالتواء Skewness measures

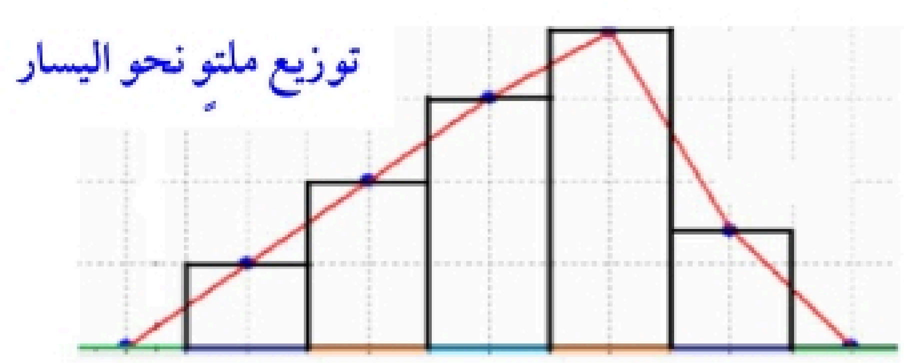
من المعلوم (وكما لاحظنا لدى بعض التوزيعات التكرارية) أنَّ التوزيع قد يكون له أحد الأشكال الثلاثة الآتية:

أ – ملتوٍ نحو اليمين، أي له ذيل منبسط نحو اليمين أكثر منه نحو اليسار.



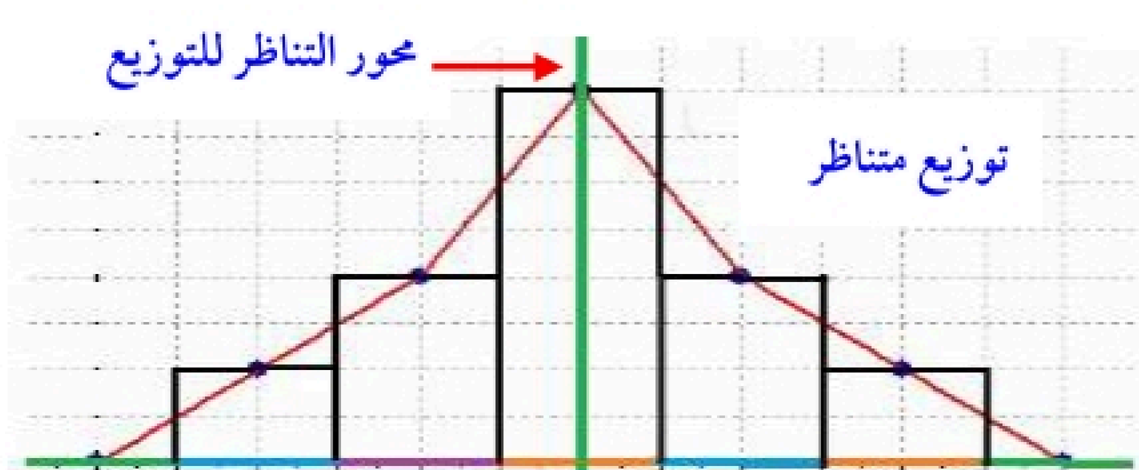

الشكل (١١ , ٢ . أ)

ب– ملتوٍ نحو اليسار، أي له ذيل منبسط نحو اليسار أكثر منه نحو اليمين.




الشكل (١١ , ٢ . ب)

ج– متناظر، أي الجزء الأيمن من التوزيع (بدءًا من قمَّة التوزيع) هو صورة مرآة لجزئه الأيسر.




الشكل (١١ , ٢ . ج)

لذلك سنقدِّم فيما يلي بعض المقاييس التي تحدِّد لنا إن كان التوزيع متناظراً أم ملتوياً، وفي حال كان ملتوياً نعيِّن مقدار هذا الالتواء وجهته. إنَّ أشهر وأدق المقاييس الخاصّة بالالتواء هو مُعامل الالتواء العزومي الذي يقدّمه لنا التعريف الآتي.

### (٢ , ٥ , ٢ , ١) تعريف (مُعامل الالتواء العزومي Moment Coefficient of Skewness):

لنفترض أنَّ لدينا بيانات مُعطاة، فعندئذ:

أ –إذا كانت هذه البيانات لعيِّنة، وكان انحرافهاالمعيّاري يساوي $S$، فإنَّ **مُعامل الالتواء العزومي** لهذه البيانات سنرمز له بـ $\mathcal{S}$، ويحسب بالعلاقة الآتية:

$$\mathcal{S} = \frac{\mathbf{S}^{(3)}}{S^3} \qquad [2\text{–}41\text{-a}]$$

فنلاحظ هنا أنَّه يجب أن يكون $S > 0$ ، وهذا محقَّق في جميع الحالات التي يكون فيها حجم العيِّنة أكبر أو يساوي 2، ومنهما قيمتان على الأقل مختلفتان عن بعضهما.

ب- إذا كانت هذه البيانات لمجتمع، وكان انحرافها المعيّاري يساوي $\sigma$ ، فإنَّ **مُعامل الالتواء العزومي** لهذه البيانات سنرمز له بـ $\mathcal{S}$ أيضاً، ويحسب بالعلاقة الآتية:

$$\mathcal{S} = \frac{\sigma^{(3)}}{\sigma^3} \qquad [2,41\text{-b}]$$

وفي كلا الحالتين السّابقتين تُناقش قيمة $\mathcal{S}$ على النحو الآتي:

١- إذا كان $\mathcal{S} > 0$ فعندئذ يكون توزيع البيانات ملتوياً نحو اليمين، أو يُقال إنَّ توزيع البيانات موجب الالتواء.

٢- إذا كان $\mathcal{S} < 0$ فعندئذ يكون توزيع البيانات ملتوياً نحو اليسار، أو يُقال إنَّ توزيع البيانات سالب الالتواء.

٣- إذا كان $\mathcal{S} = 0$ فعندئذ يكون توزيع البيانات متناظراً، أو يُقال إنَّ توزيع البيانات متماثل.

### (٢,٥,٢,١,١) ملاحظات

١- إذا كانت البيانات المُعطاة لعيِّنة، فعندئذ يمكن استخدام العلاقة الآتية أيضاً لحساب مُعامل الالتواء:

$$\mathcal{S} = \frac{\sqrt{n\,(n-1)}}{n-2} \cdot \frac{\mathbf{S}^{(3)}}{S^3} \qquad [2,41\text{-c}]$$

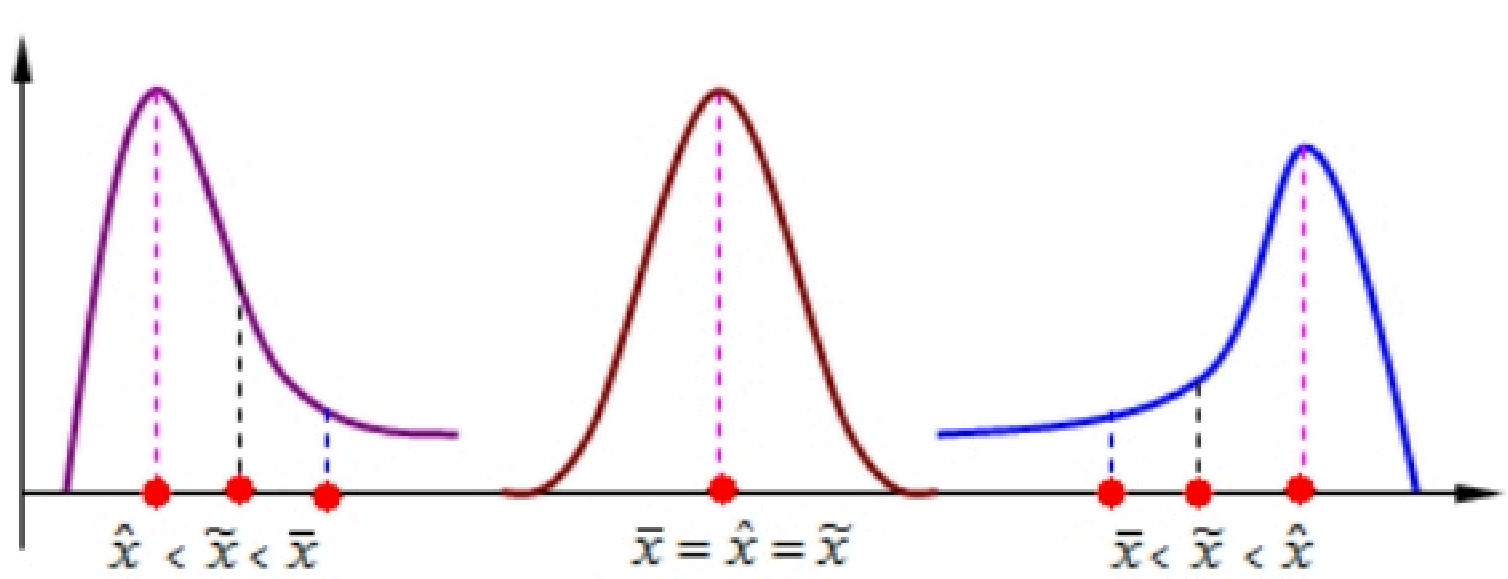
ولكن تبيَّن أنَّه من الأفضل استخدام العلاقة الآتية لحساب مُعامل الالتواء العزومي:

$$\mathcal{S} = \frac{n^2}{(n-1)\cdot(n-2)} \cdot \frac{\mathbf{S}^{(3)}}{S^3} \qquad [2,41\text{-d}]$$

لاحظ هنا يجب أن يكون $n \geq 3$ ، وهذا يجعل قيمة $S > 0$ مُحقَّقة. إنَّ الصيغة الأخيرة [2–41-d] تستخدم في بعض البرامج الإحصائية (مثل: SPSS، Minitab و...) لحساب مُعامل الالتواء من أجل بيانات العينات والمجتمعات على حد سواء، بمعنى أنَّه من أجل هذه الحالة تقوم بعض البرامج بالتعامل مع البيانات على أنَّها بيانات عيِّنة.

### (٢,٥,٢,٢) مُعامل بيرسون للالتواء Person's Coefficient of Skewness

لو أمعنا النظر في العلاقات المختلفة لحساب مُعامل الالتواء العزومي فإنَّنا نلاحظ أنَّها تتطلب تنفيذ عمليات حسابية مطوَّلة للحصول على النتيجة، ولهذا سنقدم علاقات أكثر بساطةً (ولو أنَّها أقل دقةً) لحساب قيمة مُعامل الالتواء. إنَّ الفكرة لهذه المقاييس التي سنقدمها مبنية على الوضع النسبي لقيم المتوسط، والوسيط، والمنوال لدى حالات مختلفة من الالتواء، توضحها لنا الأشكال الآتية (من أجل توزيعات مستمرة):



الشكل (١٢ , ٢)

وكما نلاحظ من الشكل السابق أنّه عندما يكون التوزيع ملتو فإنَّ المتوسط سوف ينزاح باتجاه ذيل التوزيع أكثر من الوسيط، والوسيط سينزاح باتجاه ذيل التوزيع أكثر من المنوال (علماً أنَّ قيمة المنوال توافق قمة التوزيع دوماً)، ومن ثَم فمن المناسب أن تُستخدم الفروق بين هذه القيم (المتوسط، والوسيط، والمنوال) لصنع مقياس للالتواء، ولكن مقدار هذا الفرق يعتمد على مقدار تشتّت (أو انتشار) البيانات أيضاً؛ ولذا فمن الضروري أن نتخلص من تأثير هذا التشتّت لكي نحصل على مقياس دقيق للالتواء، ويتم ذلك بقسمة ناتج الفرق على قيمة الانحراف المعّياري.

لقد قام الإحصائي الإنكليزي بيرسون Karl Pearson (1875-1936) (كان رائداً في الإحصاء ومبتكراً لاختبار كاي مربع) باستخدام هذه الميزات المذكورة آنفًا لصياغة ما يُعرف باسم مُعامل بيرسون للالتواء الذي قدَّمه على النحو الآتي:

بفرض أنَّه لدينا بيانات عيِّنة مُعطاة بمتوسط $\overline{x}$ ومنوال $\hat{x}$ وانحراف معياري $S$، فعندئذ يُحسب مُعامل الالتواء لهذه البيانات بالعلاقة الآتية:

$$\mathcal{S}_P = \frac{\overline{x} - \hat{x}}{S} \qquad [2,42\text{-a}]$$

وتُعرف هذه القيمة باسم مُعامل بيرسون للالتواء، وتناقش كما سبق لدى مُعامل الالتواء العزومي، ولكن نعلم أنَّه قد يتعذر علينا حساب المنوال $\hat{x}$ في بعض الحالات، وفي حالات أخرى قد يكون للبيانات أكثر من منوال، فعندئذ العلاقة السّابقة لا تفيدنا بشيء، ولهذا السبب يُلجأ إلى استخدام العلاقة الآتية كبديل للعلاقة السّابقة في حساب مُعامل الالتواء:

$$\mathcal{S}_P = \frac{3(\overline{x} - \tilde{x})}{S} \qquad [2,42\text{-b}]$$

وتناقش هذه القيمة $\mathcal{S}_P$ وفقاً لما سبق وتحمل اسم العلاقة السّابقة نفسه أيضاً.

### (١ , ٢ , ٢ , ٥ , ٢) ملاحظات

١- إذا كانت البيانات لمجتمع إحصائي بمتوسط $\mu$، منوال $\hat{\mu}$، وسيط $\tilde{\mu}$ وانحراف معياري $\sigma$، فعندئذ يصبح للعلاقتين السابقتين [2–42-a] و [2–42-a] العرضين الآتيين على الترتيب:

$$\mathcal{S}_P = \frac{\mu - \hat{\mu}}{\sigma} \qquad [2,42\text{-c}]$$

$$\mathcal{S}_P = \frac{3(\mu - \tilde{\mu})}{S} \qquad [2,42\text{-d}]$$

٢- إنَّ القيمة التي نحصل عليها من هذا المقياس هي قيمة عددية مجرَّدة من وحدة القياس.

### (٣ , ٢ , ٢ , ٥ , ٢) مُعامل باولي للالتواء Bowley's Coefficient of Skewness:

لو أمعنا النظر في العلاقات السّابقة لحساب مُعامل الالتواء فإنَّنا نلاحظ أنَّها تصبح عديمة التطبيق في حال فقدان بعض البيانات (ومن ثَمَّ عدم إمكانية حساب العزوم المركزيَّة للبيانات) أو وجود قيم متطرِّفة تؤدي إلى إعطاء تقدير غير مقبول للمتوسط الحسابي $\overline{x}$، ولهذا يجب البحث في علاقات بديلة تُمكّننا من تعيين الالتواء وحساب مقداره في مثل هذه الحالات.

في الواقع يمكن اللجوء إلى استخدام الرُبَيعيّات في مثل هذه الحالات إذا كانت البيانات المفقودة لا تؤثر في حساب الرُبَيعيّات، وهذا ما قام به الإحصائي والاقتصادي البريطاني باولي Sir Arthur Lyon Bowley (1869-1957) (وكان رائداً في استخدام تقنيات أخذ العينات في المسوح الاجتماعية) حيث قدّم ما يُعرف باسم مُعامل باولي للالتواء، وذلك على النحو الآتي:

بفرض أنّه لدينا بيانات مُعطاة رُبيعياته $Q_1$ ، $Q_2$ و $Q_3$ ، فعندئذ يُمكن حساب مُعامل الالتواء لهذه البيانات باستخدام العلاقة الآتية:

$$\mathcal{S}_B := \frac{(Q_3 - Q_2) - (Q_2 - Q_1)}{Q_3 - Q_1} = \frac{Q_3 + Q_1 - 2Q_2}{Q_3 - Q_1} \qquad [2,43]$$

وتُعرف هذه القيمة باسم **مُعامل باولي للالتواء** وتُناقش قيمته $\mathcal{S}_B$ وفقاً لما سبق أيضاً.

إنَّ العلاقة السّابقة تفسَّر هندسياً على النحو الآتي:

نعلم أنَّ مساحة المدرج التّكراري بين الرُبَيعيّ الأول والثاني تساوي ربع المساحة الكلية، وبالمثل مساحة المدرج التّكراري بين الرُبَيعيّ الثاني والثالث هي ربع المساحة الكلية أيضاً، ومن ثَم ستكون من أجل البيانات ذات الالتواء الموجب المسافة بين الرُبَيعيّ الثالث والثاني أقصر من المسافة بين الرُبَيعيّ الأول والثاني ، وعلى العكس في حالة الالتواء نحو اليسار، أي إن المسافة بين الرُبَيعيّ الثالث والثاني أقصر من المسافة بين الرُبَيعيّ الأول والثاني (انظر الشكل الآتي):


الشكل (١٣ , ٢)

ومن ثَمَّ يمكن استخدام الفرق $(Q_3 - Q_2) - (Q_2 - Q_1)$ للتعبير عن الالتواء، ومن أجل إلغاء تأثير تشتّت البيانات في قيمة الالتواء والتخلص من وحدة القياس بآن واحد قام بالتقسيم على المدى الرُبَيعيّ $Q_3 - Q_1$ ليحصل بذلك على العلاقة [2,43].

**(٢ , ٥ , ٢ , ٤) مثال**

١- أخذت عيِّنة من الحمام الزاجل ورُقبت أعمارها، فكانت النتائج كما في مجوعة البيانات الآتية (مقدراً بالسنة):

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 9 | 4 | 11 | 11 | 8 | 14 | 13 | 4 | 6 | 9 |
| 4 | 8 | 9 | 15 | 12 | 7 | 3 | 7 | 10 | 5 | 12 |
| 3 | 4 | 1 | 8 | 8 | 6 | 6 | 2 | 8 | 10 | |

وبعد ترتيبها تصاعدياً يصبح لها العرض الآتي:

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 2 | 2 | 3 | 3 | 4 | 4 | 4 | 4 | 5 | 6 |
| 6 | 6 | 7 | 7 | 8 | 8 | 8 | 8 | 8 | 9 | 9 |
| 9 | 10 | 10 | 11 | 11 | 12 | 12 | 13 | 14 | 15 | |

فنجد أنَّ قيمة المتوسط والوسيط لهذه البيانات هي على الترتيب:

$$\overline{x} = \frac{1}{32} \sum_{i=1}^{32} x_i = \frac{239}{32} = 7.469$$

$$\tilde{x} = \frac{x_{\frac{n}{2}} + x_{\frac{n}{2}+1}}{2} = \frac{x_{16} + x_{17}}{2} = \frac{8+8}{2} = 8 = \mathbf{Q}_2$$

$$\hat{x} = 8$$

وأَمّا قيمتي الرُّبَيْعيّ الأول والثالث فهما:

$$\mathbf{Q}_1 = x_8 + 0.25(x_9 - x_8) = 4 + 0.25(4-4) = 4$$

$$\mathbf{Q}_3 = x_{24} + 0.75(x_{25} - x_{24}) = 10 + 0.75(10-10) = 10$$

وذلك لأنَّه لدينا:

$$q_1 = \frac{n+1}{4} = \frac{32+1}{4} = 8.25 \quad \& \quad q_3 = \frac{3(n+1)}{4} = \frac{99}{4} = 24.75$$

وأمّا التباين لهذه البيانات فنجده يساوي:

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2 = \frac{419.9688}{31} = 13.547$$

ومن ثَمَّ تكون لدينا قيمة الانحراف المعيّاري $S = 3.681$. أخيراً نجد أنَّ العزم المركزي من المرتبة الثالثة لهذه البيانات يساوي:

$$\mathbf{S}^{(3)} = \frac{1}{n-1} \sum_{i=1}^{n} \left(x_i - \overline{x}\right)^3 = \frac{224.8081}{31} = 7.252$$

وهكذا يكون لُعامل الالتواء العزومي للبيانات المُعطاة القيمة الآتية:

$$\mathcal{S} = \frac{\mathbf{S}^{(3)}}{S^3} = \frac{7.252}{49.877} = 0.144$$

أو باستخدام العلاقة [2,41, d] (الصيغة المعدّلة لمعامل الالتواء) نجد لُعامل الالتواء العزومي للبيانات المُعطاة القيمة الآتية:

$$\mathcal{S} = \frac{n^2}{(n-1)\cdot(n-2)} \cdot \frac{\mathbf{S}^{(3)}}{S^3} = \frac{1024}{930} \times 0.144 = 0.160$$

فنجد من الصيغتين السّابقتين لمعامل الالتواء العزومي أنَّ توزيع البيانات ملتوٍ نحو اليمين بشكل طفيف. أمّا إذا استخدامنا مُعامل بيرسون للالتواء (العلاقة [2,42-a]) فإنَّنا سنجد:

$$\mathcal{S}_P = \frac{\overline{x} - \hat{x}}{S} = \frac{7.469 - 8}{3.681} = -0.1443$$

أو باستخدام العلاقة [2,42, b] فنجد:

$$\mathcal{S}_P = \frac{3(\overline{x} - \tilde{x})}{S} = \frac{3(7.469 - 8)}{3.681} = -0.433$$

وأخيراً إذا استخدامنا مُعامل باولي للالتواء يكون لدينا:

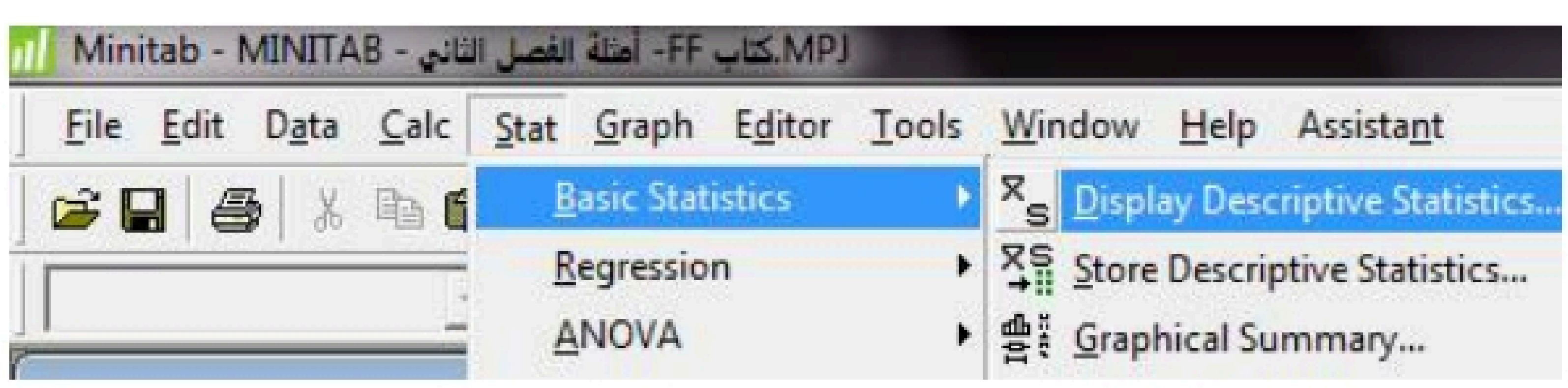
$$\mathcal{S}_B = \frac{\mathbf{Q}_3 + \mathbf{Q}_1 - 2\mathbf{Q}_2}{\mathbf{Q}_3 - \mathbf{Q}_1} = \frac{10 + 4 - 16}{10 - 4} = -0.3333$$

وهذه الصيغ الثلاث الأخيرة تُظهر لنا أنَّ توزيع البيانات ملتو نحو اليسار بشكل طفيف، وبما أنَّ هذه النتائج متناقضة مع نتائج مُعاملات الالتواء العزومية فإنَّنا نأخذ بقيم مُعاملات الالتواء العزوميّة لأنَّها الأكثر دقَّة. وهكذا نخلص إلى الملاحظات الآتية.

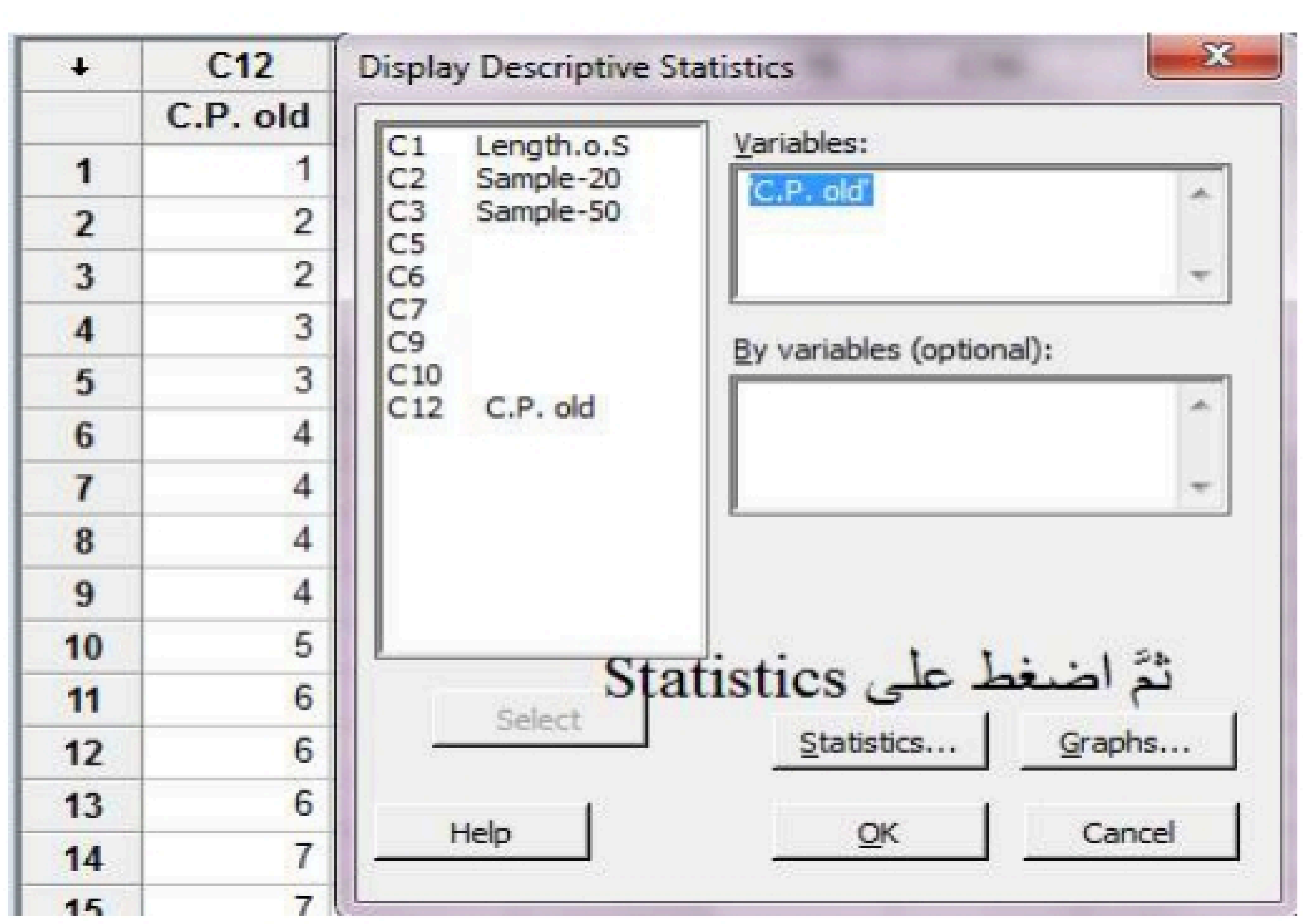
### (٥,٢,٥,٢) ملاحظات

١- بما أنَّ مقاييس الالتواء لـ **بيرسون** و **باولي** ليست دقيقة بما يكفي إذا ما قورنت بقيمة مُعامل الالتواء العزومي، فإنَّه إذا قمنا باستخدام أحد هذه المقاييس (مقاييس بيرسون و باولي) ولاحظنا أنَّ التوزيع متناظر أو ملتو بشكل طفيف جداً فإنَّه من الأفضل (بل الواجب في حال الضرورة) اللجوء إلى استخدام مُعامل الالتواء العزومي لإظهار حقيقة التناظّر أو الالتواء. بمعنى أنَّه أدقّ القيم لمُعامل الالتواء هي تلك التي تُحسب باستخدام مُعامل الالتواء العزومي، ويليها مُعامل بيرسون للالتواء، ومن ثمَّ مُعامل باولي للالتواء.

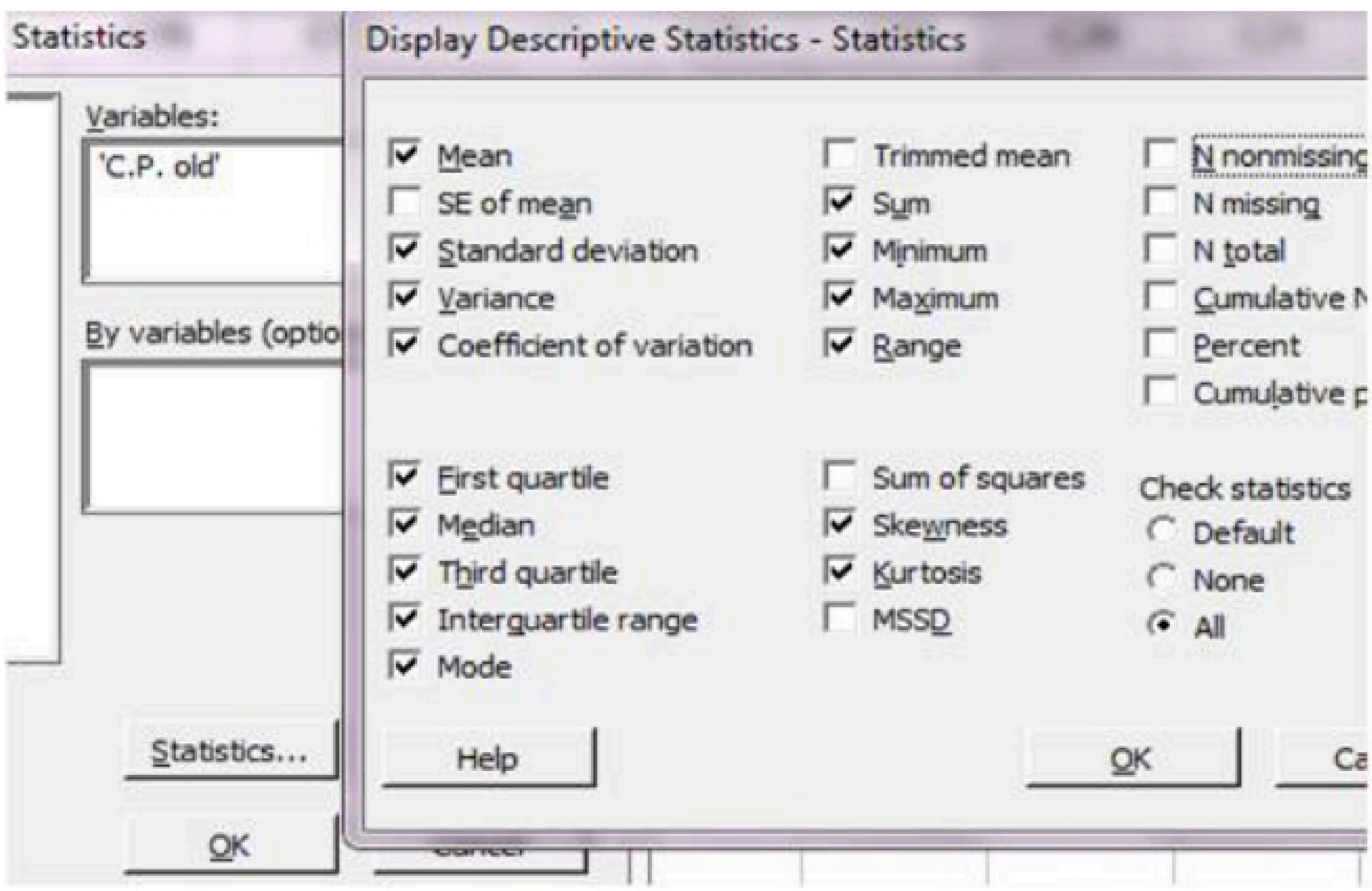
٢- من الممكن استخدام برنامج Minitab للحصول على النتائج السّابقة، وذلك وفقاً للخطوات الموضَّحة بالعروض الآتية:




ومن ثمَّ نضغط على Statistics




وبعد ذلك فعِّل النوافذ التي ترغب بها في اللوحة التالية ومن ثمَّ اضغط OK ثمَّ OK على اللوحة السّابقة:




فتحصل على التقرير الآتي (قمنا بالتعديل على التنسيق فقط لإظهاره بشكل واضح):

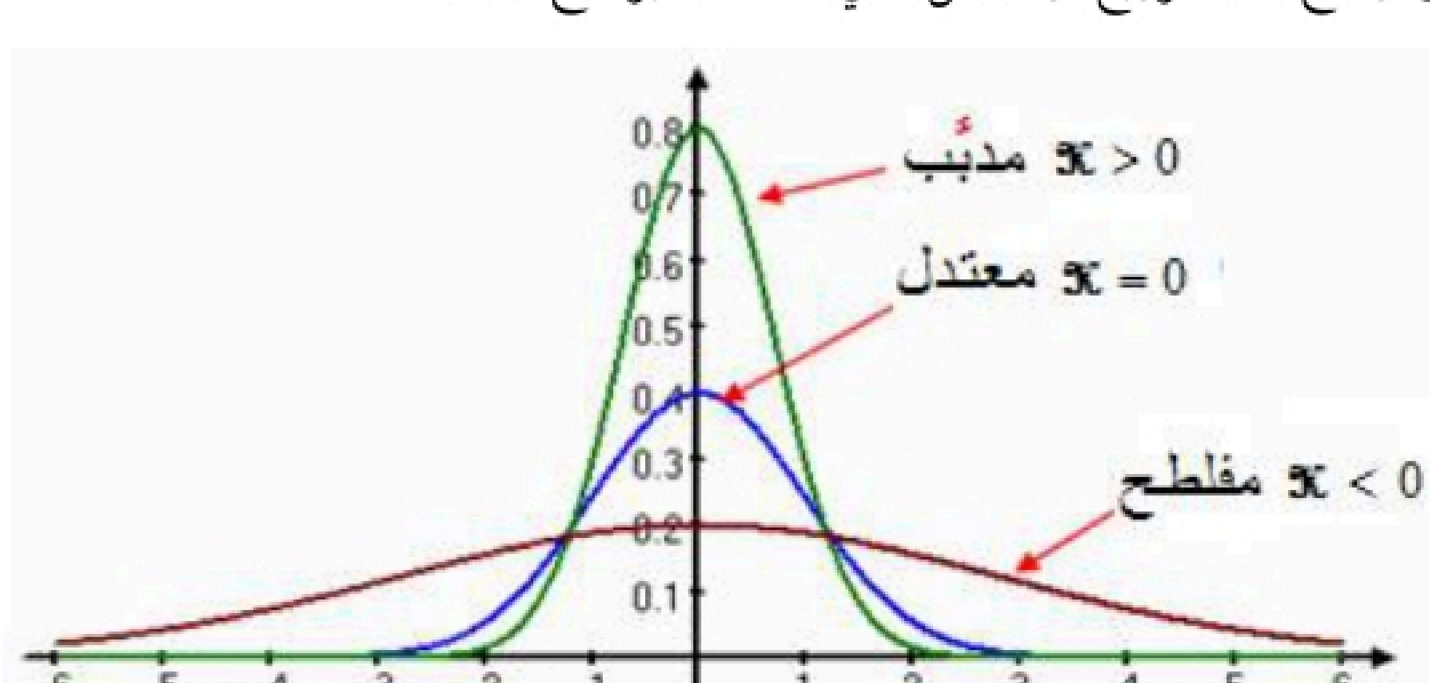
**Descriptive Statistics:**

| Mean | StDev | Variance | CoefVar | Sum | Minimum | $Q_1$ | Median | $Q_3$ | Maximum |
|---|---|---|---|---|---|---|---|---|---|
| 7.469 | 3.681 | 13.547 | 49.28 | 239.0 | 1.0 | 4.0 | 8.0 | 10.0 | 15.0 |

| Range | IQR | Mode | Mode | Skewness | Kurtosis |
|---|---|---|---|---|---|
| 14.0 | 6.0 | 8 | 5 | 0.16 | -0.71 |

### (٢,٥,٣) مقاييس التفلطح Kurtosis Measures

من المعلوم أنَّه من أجل بيانات مُحدَّدة ستكون المساحة تحت الشكل الممثَّل للتوزيع (مدرج تكراري، أو مضلَّع تكراري، أو منحنى تكراري) تساوي مقدارًا ثابتًا، ولهذا السبب، فإنَّه إذا كان تشتّت البيانات كبيراً فإنَّ ذلك سيوافق امتداداً لأحد ذيلي التوزيع (على الأقل)، ومن ثَمَّ انخفاضاً لقمَّة التوزيع. وعلى العكس، فإذا كان تشتّت البيانات صغيراً فإنَّ ذيلي التوزيع سوف ينبسطان مقتربين من المتوسط، ومن ثَمَّ سيؤدي ذلك إلى ارتفاع قمة التوزيع، والشكل الآتي (٢,١٤) يوضِّح ذلك.




الشكل (٢,١٤)

الآن من أجل قياس درجة هذا الانخفاض أو الارتفاع لقمة التوزيع وما يرافق ذلك لأطراف التوزيـع اعتمدَت مقاييس عديدة سُـمِّيت بمقاييس **التفلطح** (أو مقاييس التفرطح)، علماً أنَّ تفلطح توزيع ما يُقاس بمقدار **انبساط** أو **تدبب** المنحَنى التّكراري لتوزيع البيانات، ولهذا كان لا بدَّ من إعطاء مقاييس عددية تُحدّد مقدار هذا الانبساط أو التدبّب. في هذا الصدد اقترح **بيرسون** العلاقة الآتية لقياس درجة تفلطح توزيعٍ ما.

### (١, ٣, ٥, ٢) مُعامل بيرسون للتفلطح Person's Coefficient of Kurtosis

لنفترض أنَّ لدينا بيانات مُعطاة لعيِّنة انحرافها المعيّاري $S$، فعندئذ يُدعى المقدار الآتي:

$$\mathcal{K} := \frac{\mathbf{S}^{(4)}}{S^4} - 3 \qquad [2,44\text{-a}]$$

**مُعامل التفلطح العزومي** للبيانات.

أمّا إذا كانت البيانات المُعطاة لمجتمع بانحراف معياري $\sigma$، فعندئذ يُحسب مُعامل التفلطح العزومي لهذه البيانات من خلال العلاقة الآتية:

$$\mathcal{K} := \frac{\sigma^{(4)}}{\sigma^4} - 3 \qquad [2,44\text{-b}]$$

إنَّ العلاقتين السابقتين [2,44-a] و [2,44-b] هما الأصل في تعريف مُعامل التفلطح العزومي، وقيمة هذا المُعامل $\mathcal{K}$ تُناقش على النحو الآتي:

١- إذا كان $\mathcal{K} > 0$، فعندئذ يكون التوزيع **مُدبَّباً**، ويزداد تدبُّباً كلما ازدادت قيمة $\mathcal{K}$.

٢- إذا كان $\mathcal{K} < 0$، فعندئذ يكون التوزيع **منبسطاً**، ويزداد انبساطاً كلما تناقصت قيمة $\mathcal{K}$.

٣- إذا كان $\mathcal{K} = 0$، فعندئذ يكون التوزيع**معتدلاً**، علماً أنَّ القيمة (0) أخذت كمقياس للاعتدال، وذلك على افتراض أنَّ قيمة $\mathcal{K}$ للتوزيع الطبيعي المعّياري (سنأتي على ذكره في الفصل السادس) تساوي الصفر وفقاً للعلاقة السّابقة، والنظر إلى شكل التوزيع الطبيعي المعّياري كنموذجٍ للتوزيعات معتدلة التفلطح كما ذكرنا ذلك في الفصل السّابق (انظر الشكل السّابق (١٤ , ٢) لدالّة الكثافة الاحتمالية للتوزيع الطبيعي المعيّاري من أجل قيم مختلفة لـ $\mathcal{K}$).

### (١, ١, ٣, ٥, ٢) ملاحظات

١- إذا كانت البيانات المُعطاة لعيِّنة، فعندئذ يُفضَّل استخدام الصيغة المُعدَّلة الآتية لحساب مُعامل التفلطح:

$$\mathcal{K} := \frac{n^2(n+1)\cdot \mathbf{S}^{(4)}}{(n-1)(n-2)(n-3)\cdot S^4} - 3\frac{(n-1)^2}{(n-2)\cdot(n-3)} \qquad [2,44\text{-c}]$$

وتناقش قيمة $\mathcal{K}$ كما سبق.

٢- إنَّ العلاقة الأخيرة [2,44-c] تُستخدم في بعض البرامج الإحصائية (مثل: **Excel**، **SPSS**، **Minitab** و...) لحساب مُعامل الالتواء من أجل بيانات العينات والمجتمعات على حدٍّ سواء، بمعنى أنَّه من أجل هذه الحالة تقوم بعض البرامج بالتعامل مع البيانات على أنَّها بيانات عيِّنة.

**(٢ , ١ , ٣ , ٥ , ٢) أمثلة**

١- بالرجوع إلى المثال السّابق (٤ , ٢ , ٥ , ٢) حيث لدينا $S = 3.681$، وبحساب العزم المركزي من المرتبة الرابعة لتوزيع تلك البيانات نجده يساوي:

$$\mathbf{S}^{(4)} = \frac{1}{31} \sum_{i=1}^{32} \left(x_i - \overline{x}\right)^4 = \frac{12193.45}{31} = 393.3371$$

ومن ثمَّ تكون قيمة مُعامل التفلطح لتوزيع البيانات المعطاة يساوي:

$$\mathcal{K} = \frac{\mathbf{S}^{(4)}}{S^4} - 3 = \frac{393.3371}{183.596} - 3 = -0.8576$$

أو وفقاً للصيغة المُعدَّلة لمُعامل التفلطح:

$$\mathcal{K} = \frac{n^2 \cdot (n+1)}{(n-1)\cdot(n-2)\cdot(n-3)} \cdot \frac{\mathbf{S}^{(4)}}{S^4} - 3\frac{(n-1)^2}{(n-2)\cdot(n-3)} = \frac{33792}{26970} \cdot \frac{393.3371}{183.596} - 3\frac{1024}{870} = -0.71$$

وهي النتيجة ذاتها التي حصلنا عليها باستخدام برنامج **Minitab** (انظر التقرير السّابق) التي تبيّن لنا أنَّ توزيع البيانات المُعطاة منبسط.

## (٢,٦) الأعداد الخمسة والتمثيل الصندوقي للبيانات

**Five Number and the Box plot**

يفي الواقع تستخدم المخططات الصندوقية في معظم التحاليل الإحصائية الوصفية لأنَّها تعطي انطباعاً سريعاً للقارء حول قيم الُربيعيَّات الثلاثة وكذلك حول الالتواء الخاص بتوزيع البيانات.

**(٢ , ٦ , ١) تعريف (الأعداد الخمسة لبيانات كمّيّة خام)**

بفرض أنَّه لدينا $x_1$، $x_2$، ...و $x_n$ بيانات خام مُعطاة. عندئذ يُقصد بالأعداد الخمسة (في الإحصاء الوصفي) القيم الآتية:

- **أصغر قيمة** في البيانات $x_s$،
- **أكبر قيمة** في البيانات $x_\ell$،
- **قيم الُرّبيعيّات الثلاثة** $\mathbf{Q}_1$، $\mathbf{Q}_2$ و $\mathbf{Q}_3$.

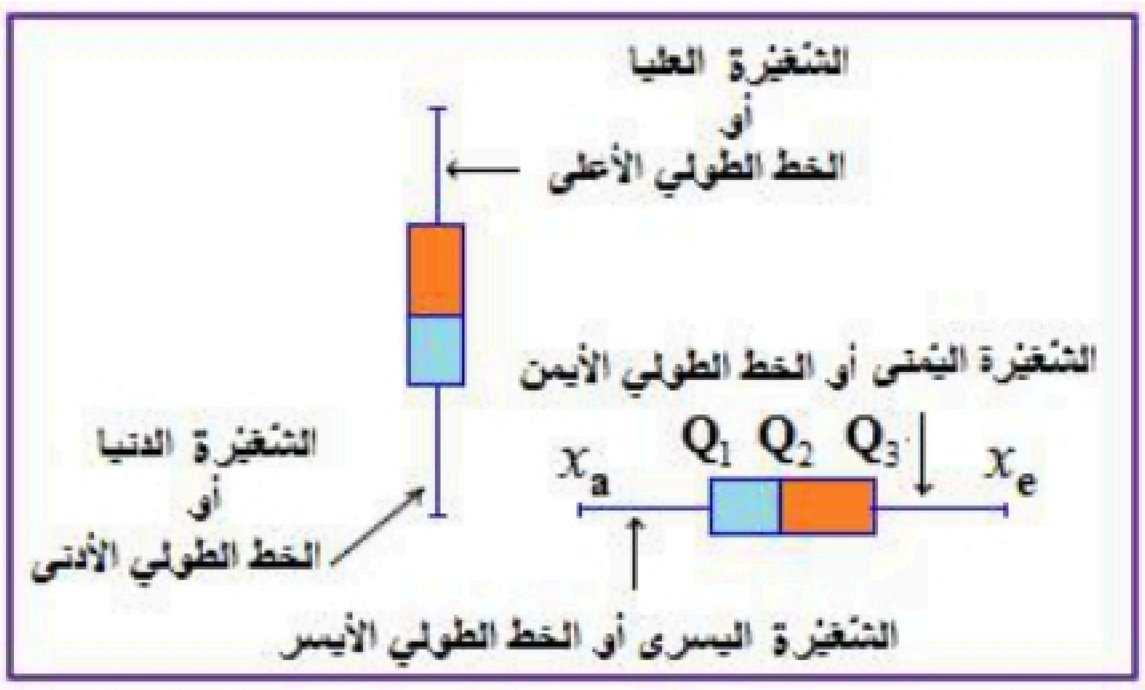
إنَّ هذه الأعداد الخمسة تساعدنا في وصف تمركز البيانات، وانتشارها، وشكل توزيعها (من أجل التوزيعات المستمرة)، والجدول الآتي يوضِّح لنا ملخّص العلاقات بين الأعداد الخمسة وشكل التوزيع للبيانات.

الجدول (٢ , ١٠)

| | | | |
|---|---|---|---|
| **Left skewed** ملتوٍ نحو اليسار | $\mathbf{Q}_2 - x_s > x_\ell - \mathbf{Q}_2$ | $\mathbf{Q}_1 - x_s > x_\ell - \mathbf{Q}_3$ | $\mathbf{Q}_2 - \mathbf{Q}_1 > \mathbf{Q}_3 - \mathbf{Q}_2$ |
| **Symmetric** متناظر | $\mathbf{Q}_2 - x_s = x_\ell - \mathbf{Q}_2$ | $\mathbf{Q}_1 - x_s = x_\ell - \mathbf{Q}_3$ | $\mathbf{Q}_2 - \mathbf{Q}_1 = \mathbf{Q}_3 - \mathbf{Q}_2$ |
| **Right skewed** ملتوٍ نحو اليمين | $\mathbf{Q}_2 - x_s < x_\ell - \mathbf{Q}_2$ | $\mathbf{Q}_1 - x_s < x_\ell - \mathbf{Q}_3$ | $\mathbf{Q}_2 - \mathbf{Q}_1 < \mathbf{Q}_3 - \mathbf{Q}_2$ |

### (٢,٦,٢) التمثيل الصندوقي لبيانات كمّيّة خام

إنَّ العرض (أو التمثيل) الصندوقي Box plot لبيانات خام هو عرض رسومي للبيانات يقوم على أساس مُلخَّص الأعداد الخمسة السّابقة، ويتكوَّن من صندوق وخطين طوليين يتوسطانه على طرفيه يمنة ويسرى يُدعيان "الشُّعَيْرتان Whiskers". ويُقدَّم هذا العرض وفقاً لأحد الشكلين الآتيين:




لشكل (٢,١٥)

تُعيّن نهاية الشُّعَيْرتان اليمنى $x_e$ واليسرى $x_a$ كما يلي:

١- إذا كانت مجموعة البيانات لا تحتوي قيمًا متطرّفة، فعندئذ يكون لدينا:

$$x_a = x_s \qquad \& \qquad x_e = x_\ell$$

٢- إذا كانت مجموعة البيانات تحتوي على قيم صغيرة متطرّفة فقط، فعندئذ يكون $x_a = \mathbf{LF}$. أي إنَّ نهاية الشُّعَيْرة اليسرى $x_a$ توافق الحد الأعلى للقيم المتطرّفة بصغرها، ومن ثمَّ تبلغ الشُّعَيْرة اليسرى الطول الأعظمي لها. أمّا $x_e$ فيكون مساوياً لـ $x_\ell$.

٣- إذا كانت مجموعة البيانات تحتوي على قيم كبيرة متطرّفة فقط، فعندئذ يكون لدينا $x_e = \mathbf{HF}$. أي إن نهاية الشُّعَيْرة اليمنى $x_e$ توافق الحد الأدنى للقيم المتطرّفة بكبرها، ومن ثمَّ تبلغ الشُّعَيْرة اليمنى الطول الأعظمي لها. أمّا $x_a$ فيكون مساوياً لـ $x_s$.

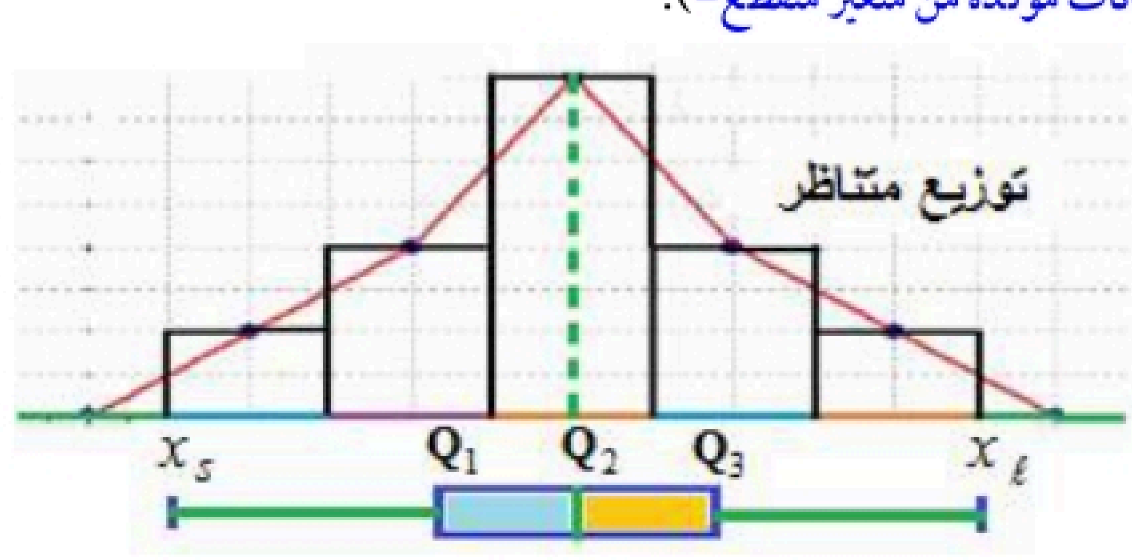
٤- إذا كانت مجموعة البيانات تحتوي على قيم صغيرة وكبيرة متطرّفة، فعندئذ يكون لدينا $x_a = \mathbf{LF}$ و $x_e = \mathbf{HF}$.

وفي هذه الحالة تبلغ الشُّعَيْرتان اليسرى واليمنى الطول الأعظمي لهما.

#### (٢,٦,١,١) ملاحظات

١- عند تقديم الرسم الصندوقي يُشار إلى كل قيمة متطرّفة بنجمة (*) أو نقطة (●).

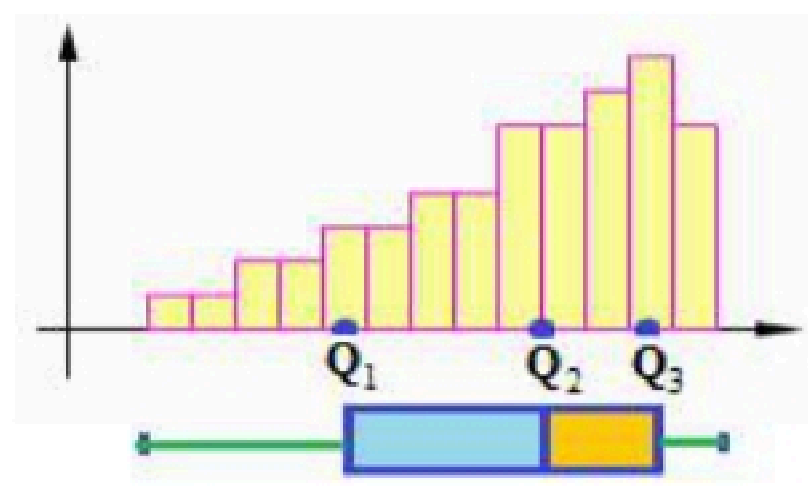
٢- من أجل البيانات المستمرة (أي أنَّها بيانات مولَّدة من متغيِّر مستمر) فإنَّه إذا كان توزيع البيانات متناظراً (أي متماثلاً حول الوسيط)، فعندئذ الصندوق وخط الوسيط سوف يتمركزان بين نقاط نهايتي الشُّعَيْرتين (انظر الشكل التوضيحي الآتي وهذه الخاصيّة تبقى صحيحة حتى من أجل البيانات المتقطّعة - أي أنَّهابيانات مولَّدة من متغيِّر متقطِّع-).




الشكل (٢,١٦.أ)

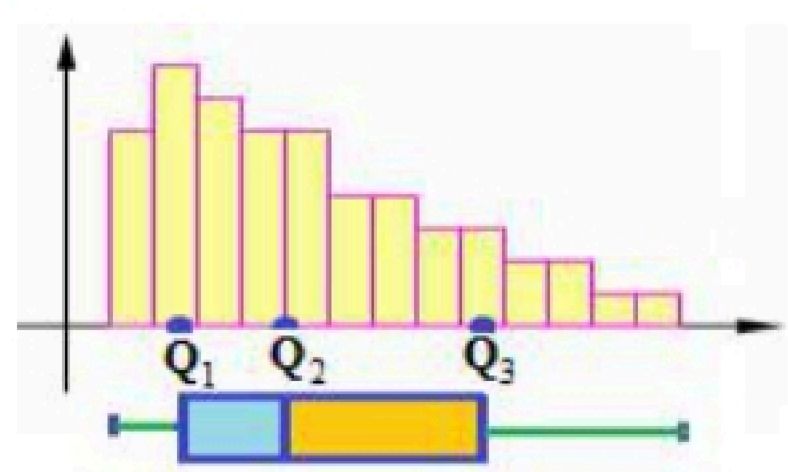
وأمّا إذا كان توزيع البيانات ملتوياً فعندئذ:

أ- من أجل التوزيعات الملتوية نحو اليسار سوف ينزاح الصندوق لليمين باتجاه نهاية الشُّعَيْرة اليمنى وأمّا خط الوسيط فيمكن أن يأخذ أية وضعية داخل الصندوق وذلك بحسب قيمته بالنسبة إلى $Q_1$ و $Q_3$ (انظر الأمثلة التوضيحية الآتية).




الشكل (١٦ , ٢ . ب)

ب- من أجل التوزيعات الملتوية نحو اليمين سوف ينزاح الصندوق لليسار باتجاه نهاية الشُعَيْرة اليسرى وأمّا خط الوسيط فيمكن أن يأخذ أية وضعية داخل الصندوق وذلك بحسب قيمته بالنسبة إلى $Q_1$ و $Q_3$ (انظر الأشكال التوضيحية الآتية).




الشكل (١٦ , ٢ . ج)

### (٢ , ١ , ٦ , ٢) أمثلة

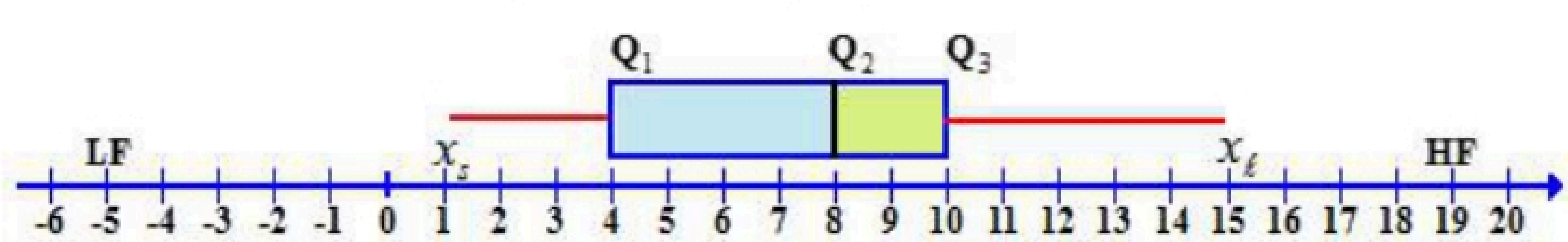
١ – بالرجوع إلى المثال السّابق (٤–٢–٥–٢) حيث وجدنا $\mathcal{S} = 0.0575$ . أي إنَّ توزيع هذه البيانات ملتوياً بشكل طفيف جداً نحو اليمين. من جهةٍ أخرى لدينا من مُعطيات المثال $x_s = 1$، $Q_1 = 4$، $\tilde{x} = Q_2 = 8$، $Q_3 = 10$ و $x_\ell = 15$، وبملاحظة أنَّ:

$$\mathbf{LF} = 4 - 1.5(10-4) = -5$$

فعندئذ القيم المتطرفة بصغرها هي تلك القيم التي تكون أصغر من 5 – ، ومن ثَّم لا توجد قيم متطرّفة بصغرها في البيانات المقدَّمة، وكذلك لدينا:

$$\mathbf{HF} = 10 + 1.5\ (10-4) = 19$$

وهذا يعني أنَّ القيم المتطرّفة بكبرها هي تلك القيم التي تكون أكبر من 19، وبالتالي ليس لدينا أية قيم متطرّفة بكبرها في البيانات المقدَّمة، ولذلك سيكون لدينا $x_{\mathbf{a}} = x_s$ وكذلك $x_{\mathbf{e}} = x_\ell$، ومن ثَّم يصبح للتمثيل الصندوقي لتلك البيانات العرض الآتي:




الشكل (١٧ , ٢ . أ)

وباستخدام برنامج **Minitab** نحصل على التمثيل الصندوقي لهذه البيانات كما في العرض الآتي:


الشكل (١٧ , ٢ . ب)

٢- لتكن لدينا مجموعة البيانات الآتية:

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **-7** | **5** | **17** | **3** | **9** | **5** | **1** | **9** | **6** | **7** | **4** | **8** |

فعندئذ نجد أنّ $x_s = -7$ ، $x_\ell = 17$ ، $\mathbf{Q}_1 = 3.5$ ، $\tilde{x} = \mathbf{Q}_2 = 5$ و $\mathbf{Q}_3 = 8.5$ ، وأَمّا لتحديد القيم المتطرّفة بصغرها التي تكون أصغر من نهاية الشُعَيْرة اليسرى، فلدينا:

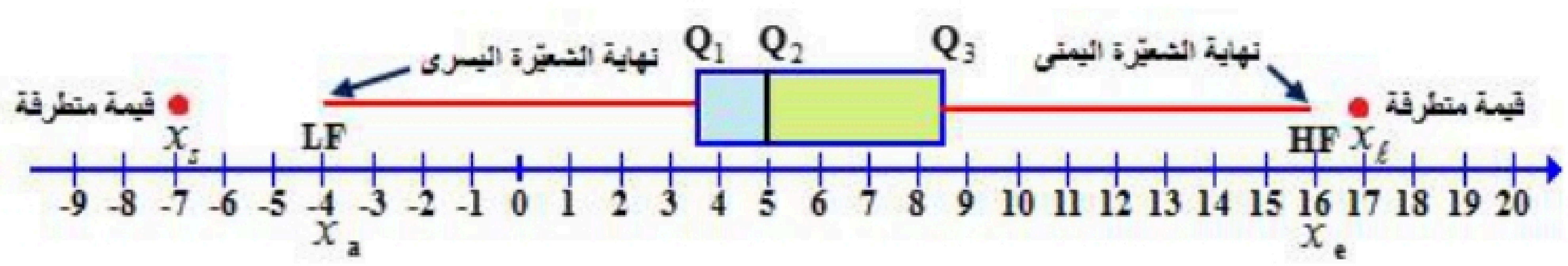
$$\mathbf{LF} = 3.5 - 1.5(8.5 - 3.5) = -4$$

أي إنَّ القيم المتطرّفة بصغرها هي تلك القيم التي تكون أصغر من 4- ، ومن ثَمَّ لدينا قيمة متطرّفة بصغرها واحدة هي $x_s = -7$ .

أخيراً من أجل تحديد القيم المتطرّفة بكبرها علينا حساب **HF** حيث لدينا:

$$\mathbf{HF} = 8.5 + 1.5(8.5 - 3.5) = 16$$

أي إنَّ القيم المتطرّفة بكبرها هي تلك القيم التي تكون أكبر من 16، وبالتالي لدينا قيمة متطرّفة بكبرها هي $x_\ell = 17$ . هكذا تبلغان الشُعَيْرتان اليمنى واليسرى الطول الأعظم لهما، حيث لدينا $x_a = \mathbf{LF} = -4$ و $x_e = \mathbf{HF} = 16$ ، ومن ثَمَّ يكون للتمثيل الصندوقي لتلك البيانات العرض الآتي:




الشكل (٢، ١٨)

## هذا ما تيسَّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل

# تمارين الفصل الثاني

١- إذا كانت $x_1$، $x_2$،...و $x_n$ مع $1 \le n$ مجموعة بيانات مُعطاة، فعندئذ:

أ- هل يمكن لقيمة عددية $x$ أن تكون مقياساً للنَّزعة المركزيَّة لهذه المجموعة من البيانات علماً أنَّها توافق $x_\ell$ (أكبر قيمة في البيانات) أو $x_s$ (أصغر قيمة في البيانات)؟

ب- بفرض أنَّ $\overline{x}$ هي قيمة المتوسط لهذه المجموعة من البيانات، فمن أجل أي قيمة لـ $n$ تصبح $\overline{x} = x_\ell$ أو $\overline{x} = x_s$ ؟

ج- من أجل أي قيمة لـ $n$ تصبح قيمة الانحراف المعيّاري لهذه المجموعة من البيانات تساوي الصفر؟

٢- لتكن لدينا مجموعة البيانات الآتية التي تمثِّل الوزن لعيّنة مكوَّنة من خمسين شخصاً بالغاً (مقدَّرةً بالكيلو غرام):

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 60 | 79 | 66 | 76 | 80 | 62 | 73 | 74 | 92 | 75 |
| 82 | 61 | 65 | 84 | 75 | 81 | 71 | 77 | 84 | 77 |
| 62 | 68 | 75 | 83 | 64 | 58 | 69 | 81 | 67 | 91 |
| 64 | 77 | 92 | 73 | 64 | 84 | 63 | 74 | 64 | 69 |
| 99 | 61 | 68 | 74 | 66 | 69 | 76 | 86 | 65 | 98 |

والمطلوب ما يلي:

أ- احسب المنوال– الوسيط– المتوسط– المتوسط الهندسي– المتوسط التوافقي- المدى- الانحراف المتوسط- التباين– الانحراف المعيّاري– مُعامل التغيُّر- مُعامل بيرسون للالتواء والتفلطح إذا علمت أنَّ البيانات المُعطاة كانت:

ب- احسب الدرجة المعيّارية $Z$ و $\mathcal{T}$ لكلّ من القيم الآتية:

58 65 74 81 99

ج- احسب مُعامل الالتواء العزومي بصيغه المختلفة، ومن ثَّم قارن هذه القيم مع قيمة مُعامل الالتواء لـ بيرسون.

د- ارسم التمثيل الصندوقي لهذه البيانات موضّحاً عليها الأعداد الخمسة.

هـ- تحت الفرض أنَّ لجميع قيم البيانات المُعطاة النصيب نفسه في الاختيار ولا يؤثِّر بعضها في البعض الآخر لدى عمليّة السحب البيانات منها، فعندئذ قم بسحب عينات عشوائية من المجتمع السّابق بحجم $n = 10$ و $n = 25$ و $n = 40$، ومن ثَّم احسب المقادير المطلوبة في البنود ( أ ) وحتى ( هـ) السّابقة، وبعد ذلك قارن بين النتائج المتقابلة. ماذا تستنتج ؟

٣- ليكن لدينا مجموعة بيانات مقدَّمة من خلال جدول توزيع تكراري له العرض الآتي:

| رقم الفئة | الحدود الفعليّة للفئة | مركز الفئة $x_i$ | تكرار الفئة $f_i$ | التكرار المتجمّع الصاعد للفئة $F_i$ |
|---|---|---|---|---|
| 1 | $50 \rightarrow 55$ | | 5 | |
| 2 | $55 \rightarrow 60$ | | | 13 |
| 3 | | | 12 | |
| 4 | | | | 40 |
| 5 | | | 17 | |
| 6 | | | | 70 |
| 7 | $80 \rightarrow 85$ | | 7 | |
| Total | ------------ | ---- | 48 | ----------------- |

والمطلوب ما يلي:

أ- أكمل بيانات هذا الجدول.

ب- احسب كلاً من المقادير الآتية في حالتي المجتمع والعيّنة:

المنوال- الوسيط- المتوسط- الانحراف المتوسط- الانحراف المعيّاري- مُعامل التغيُّر- مُعامل بيرسون للالتواء.

٤- إذا كان معدَّل العجز في ميزانية بلد ما خلال الأعوام 2008، 2009، 2010، 2011، 2012 و2013 على الترتيب هو 5%، 7%، 3%، 2%، 9% و 6%، فعندئذ احسب معدَّل العجز في ميزانية هذا البلد خلال الأعوام الستة.

٥- لتكن لدينا البيانات الآتية التي تمثِّل عيِّنة من نتائج نهاية الفصل لـ 60 طالباً في مقرَّر الإحصاء:

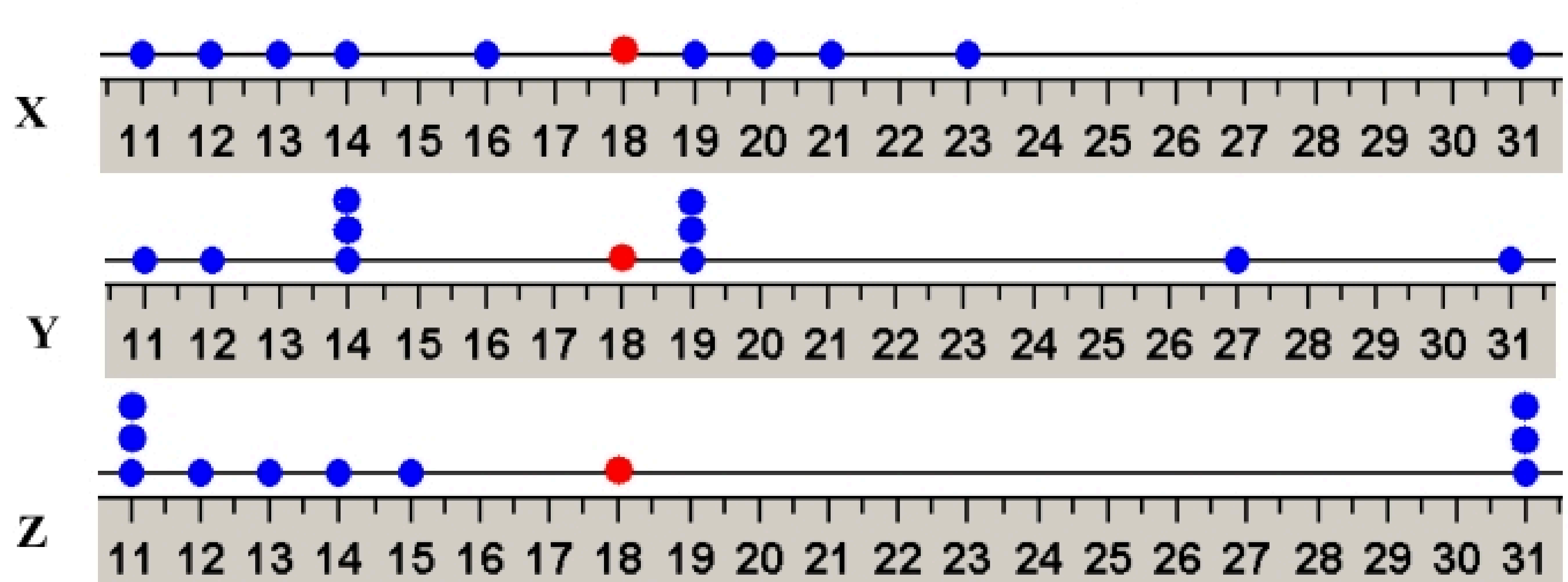
| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 52 | 99 | 92 | 86 | 84 | 52 | 99 | 92 | 86 | 84 |
| 63 | 72 | 76 | 95 | 88 | 63 | 72 | 76 | 95 | 88 |
| 92 | 58 | 65 | 79 | 80 | 92 | 58 | 65 | 79 | 80 |
| 90 | 75 | 74 | 56 | 99 | 90 | 75 | 74 | 56 | 99 |
| 65 | 79 | 80 | 90 | 75 | 74 | 56 | 99 | 90 | 75 |
| 74 | 56 | 74 | 56 | 99 | 65 | 79 | 80 | 90 | 75 |

والمطلوب ما يلي:

أ- حساب المتوسط والانحراف المعيّاري ومن ثَمَّ مُعامل الالتواء العزومي ومُعامل التفلطح لهذه البيانات.

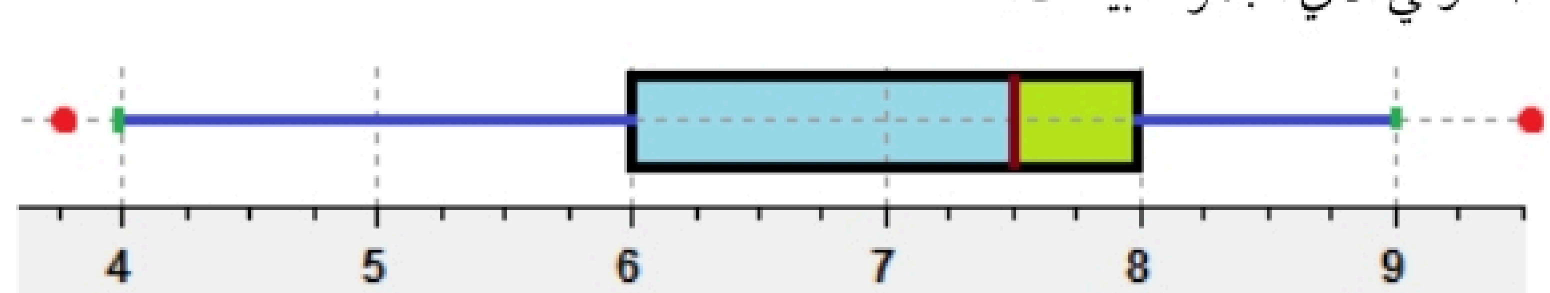
ب- قدِّر عدد القيم التي تقع في الفترة $\left[\bar{x}-2S\ ,\ \bar{x}+2S\right]$ ، ومن ثَمَّ بيِّن إن كان هذا العدد يتوافق مع قاعدة تشبيشيف التجريبية أم لا؟

٦- لتكن لدينا مجموعات البيانات الموضَّحة في العروض الآتية:




والمطلوب حساب المتوسط، والوسيط، والمدى، والانحراف المعيّاري، ومعامل التغيُّر لكل منها. ماذا تلاحظ؟ ومن ثَمَّ بيِّن أيها أقل تبعثراً من الأخرى.

٧- ليكن لدينا المخطط الصندوقي الآتي لمجموعة بيانات:




والمطلوب ما يلي:

أ- تعيين الأعداد الخمسة على هذا المخطط، $x_e$، $x_a$، LF و HF .

ب- تحديد نوع التواء التوزيع مع ذكر الأسباب.

٨- لتكن لدينا مجموعة البيانات الآتية لعيِّنة تمثِّل الوزن لمئة شخص بالغ (مقدَّرةً بالكيلو غرام):

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 84 | 63 | 74 | 64 | 69 | 69 | 76 | 86 | 65 | 98 |
| 99 | 61 | 68 | 74 | 66 | 69 | 76 | 86 | 65 | 98 |
| 64 | 77 | 92 | 73 | 64 | 84 | 63 | 74 | 64 | 69 |
| 62 | 68 | 75 | 83 | 64 | 58 | 69 | 81 | 67 | 91 |
| 64 | 77 | 92 | 73 | 64 | 84 | 63 | 74 | 64 | 69 |
| 99 | 61 | 68 | 74 | 66 | 69 | 76 | 86 | 65 | 98 |
| 64 | 77 | 92 | 73 | 64 | 84 | 63 | 74 | 64 | 69 |
| 62 | 68 | 75 | 83 | 64 | 58 | 69 | 81 | 67 | 91 |
| 64 | 77 | 92 | 73 | 64 | 84 | 63 | 74 | 64 | 69 |
| 84 | 63 | 74 | 64 | 69 | 69 | 76 | 86 | 65 | 98 |

والمطلوب ما يلي:

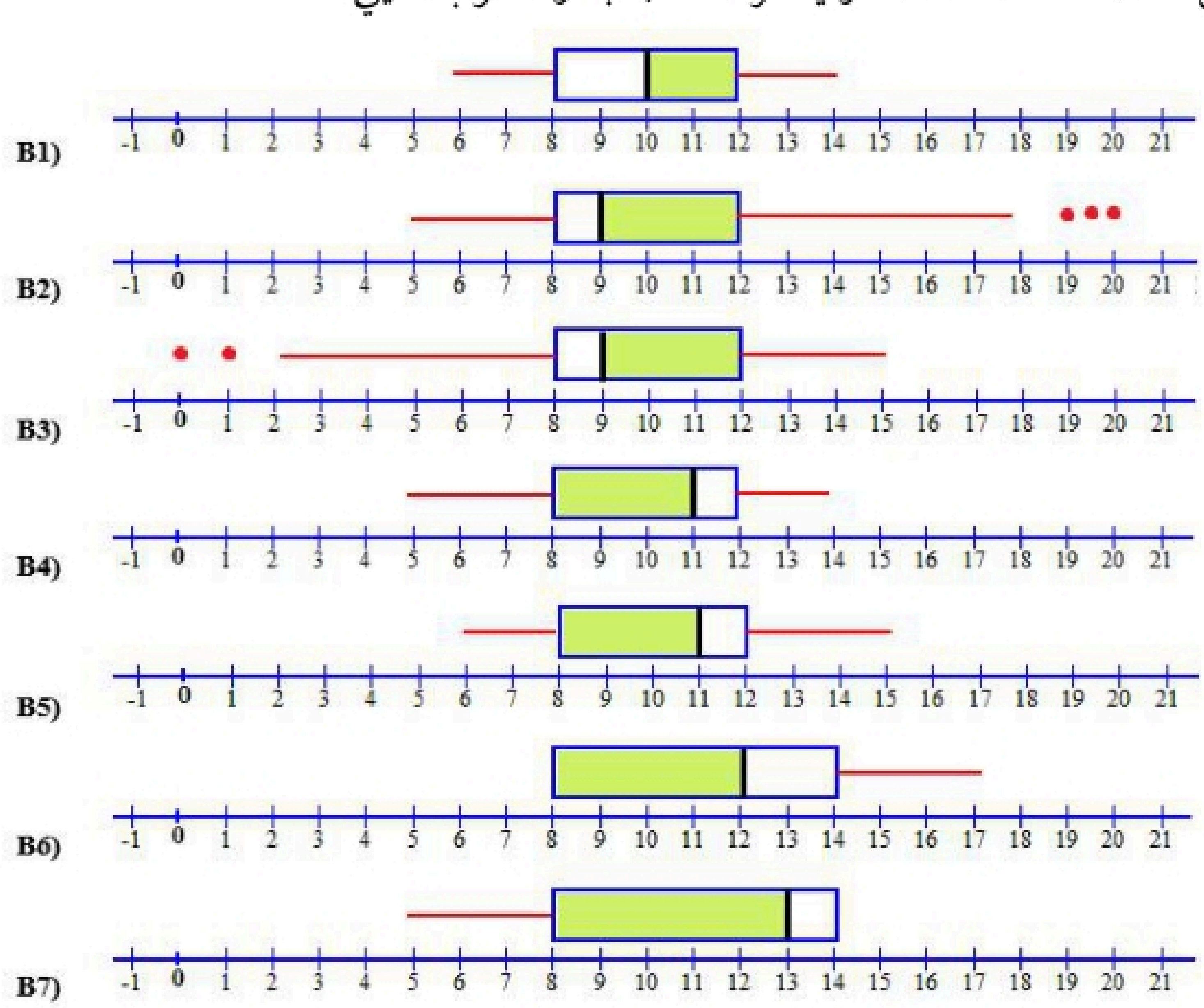
أ- احسب الرُبَيْعيّات الثلاثة لهذه البيانات.

ب - احسب المئينات $P_{35}$ ، $P_{50}$ و $P_{65}$ لهذه البيانات.

ج - ارسم المخطط الصندوقي لهذه البيانات موضِّحا عليه مواضع الأعداد الخمسة، ومن ثمَّ بيِّن نوع الالتواء لهذه البيانات.

هـ- عيِّن المئين الموافق لكل للقيم 65، 75 و98.

٩- لتكن لدينا مجموعات لبيانات مقدَّمة من خلال المخططات الصندوقية الوضَّحة جانباً، والمطلوب ما يلي:




أ- عيِّن قيم الأعداد الخمسة الخاصة بكل مخطط من المخططات (B1) وحتى (B7).

ب- أيّ الأشكال السابقة يدلّ على توزيع متناظر، ملتوٍ نحو اليمين، ملتوٍ نحو اليسار.

هـ- أيّ الأشكال السابقة يشير إلى وجود قيم متطرفة بكبرها.

و- أيّ الأشكال السابقة يشير إلى وجود قيم متطرفة بصغرها.

ز- قدِّم تفسيراً لغياب الشعيّرات لدى بعض المخططات.

# الفصل الثالث

# الارتباط وتحليل الانحدار
# CORRELATION AND REGRESSION ANALYSIS

## (٣,١) الارتباط الخطي البسيط
## Simple Linear Correlation

بعد أن قمنا في الفصلين السّابقين بدراسة مفصَّلة بعض الشيء للبيانات، حيث لاحظنا أنَّ هذه البيانات قد تكون قيماً لمتغيِّر $X$ يُعبِّر عن ظاهرة مُحدّدة $A$، وأكثر من ذلك لاحظنا أنَّه قد تكون لدينا مجموعتين أو أكثر من البيانات التي قد يكون بعضها مختلف عن بعض، مثل الطول والوزن لأشخاص. هنا قد يتساءل المرء عن التّأثير المتبادل بين ظاهرتين (أو متغيِّرين) أو أكثر، فعلى سبيل المثال: قد يتساءل المرء عن العلاقة التي تربط بين الطول والوزن لأشخاص من فئة عمريّة معيّنة (فيكون لدينا متغيِّرين)، أو بين تركيز النيكوتين في الدم، وأمراض الرئة، وتصلّب الشرايين لدى مجتمع من المدخنين (فيكون لدينا ثلاثة متغيِّرات)، أو مستوى الدخل واستهلاك بعض المواد الغذائيّة، وزيادة الوزن، واقتناء بعض أجهزة الرفاهية (فيكون لدينا أربعة متغيِّرات)، وهكذا دواليك ...

إنَّ الأدوات المستخدمة لاستكشاف مثل هذه العلاقات واستخدامها في التقدير والتنبؤ تندرج تحت مفهوم الارتباط وتحليل الانحدار، ويمكن استخدام هذه الأدوات لمعرفة هل كانت النتيجة من متغيِّر واحد $Y$ تعتمد على قيم المتغيِّرات الأخرى $X_1$، $X_2$، ... و $X_k$ أم لا؟ وهذا ما تعنيه تبعية متغيِّر لمتغيِّرات أخرى، حيث يُدعى $Y$ **متغيِّر الاستجابة** Response Variable (في الرياضيات يُدعى المتغيِّر التابع Dependent Variable) وأمّا بقية المتغيِّرات $X_1$، $X_2$، ... و $X_k$ فإنَّها تُدعى **متغيِّرات تفسيرية** Explanatory Variables (في الرياضيات تُدعى متغيِّرات مستقلّة Independent Variables).

### (٣,١,١) نماذج من الارتباط

في الواقع إنَّ مفهوم الارتباط يمكن استخدامه لوصف طبيعة العلاقة بين ظاهرتين أو أكثر، فلو قمنا برصد أكثر من ظاهرة كأن نقوم مثلاً برصد الوزن والطول لعناصر مجتمع من الأشخاص البالغين، فعندئذ يُمكننا أنَّ نُمثِّل نتيجة الدراسة لكل عنصر من هذا المجتمع بمتَّجه ذي بُعد يساوي عدد الظواهر المستخدمة في الدراسة، وبَعد ذلك يُمكننا طرح السؤال الآتي:

**هل ثمَّة علاقة ما تربط بين مُركِّبات هذا المتَّجه، وإن كانت هناك علاقة ما تربط بين هذه المركِّبات، فما هي هذه العلاقة؟**

قبل الإجابة على هذه التساؤلات سوف نتّفق على استخدام كلمة متغيِّر للدلالة على ظاهرة يمثِّلها هذا المتغيِّر (كما قدمناه في الفصلين السّابقين) الذي بدوره يقوم بتوليد البيانات الخاصّة بهذه الظاهرة، فعلى سبيل المثال: لو أخذنا الظاهرة $A$ هي ظاهرة الطول لكائنات عددها $n$، فعندئذ المتغيِّر $X$ يلعب دور أداة القياس التي تقيس الطول وتولِّد البيانات $x_1$، $x_2$، ... و $x_n$ الممثّلة لقيم الطول لهذه الكائنات. في هذه الحالة يُقال إنَّ المتغيِّر $X$ راصدٌ للظاهرة $A$ وواصف للمجتمع الإحصائي الذي تنتمي إليه هذه البيانات.

الآن للإجابة على التساؤلات السابقة سنقوم بتجزئة هذه المسألة إلى نموذجين من المسائل هي:

١ – عندما تكون المسألة متعلّقة برصد متغيِّرين $X$ و $Y$ فقط، فإنَّه في هذه الحالة يكون للبيانات شكل ثنائيّة مُرتَّبة $(x_i, y_i)$ من

أجل كل القيم الممكنة لـ $i$ ، ويُقال عن البيانات $x_i$ و $y_i$ من أجل كل القيم الممكنة لـ $i$ إنَّها **بيانات متزاوجة**، وهنا يُبحث في اتجاهين رئيسيين أيضاً هما:

أ- إذا كانت العلاقة بين البيانات المتزاوجة $x_i$ و $y_i$ هي **علاقة خطّية** من أجل كل القيم الممكنة لـ $i$ .

ب- إذا كانت العلاقة بين البيانات المتزاوجة $x_i$ و $y_i$ هي **علاقة غير خطّية** من أجل كل القيم الممكنة لـ $i$ .

٢- عندما تكون المسألة متعلّقة برصد أكثر من متغيِّرين (**ولكن عدد منته من المتغيِّرات**) $X_1$ ، $X_2$ ،... و $X_k$ . في هذه الحالة تكون البيانات الناتجة عن الدراسة مُمثَّلة بمتَّجهات $(x_{11}, x_{12}, ..., x_{1k})$ ، $(x_{21}, x_{22}, ..., x_{2k})$ ،...و $(x_{n1}, x_{n2}, ..., x_{nk})$، وهنا يُبحث في الحالتين الآتيتين أيضاً:

أ- إذا كانت العلاقة التي تربط بين متغيِّر من هذا المتَّجه $(X_1, X_2, ..., X_k)$ مع بقية مُركّباته هي **علاقة خطّية**.

ب- إذا كانت العلاقة التي تربط بين متغيِّر من هذا المتَّجه $(X_1, X_2, ..., X_k)$ مع بقية مُركّباته هي **علاقة غير خطّية**.

بالطبع سوف لن نقوم بدراسة كل هذه الفقرات من الحالات لأنَّها تحتاج إلى مجلدات لتغطيها، ولكن سنركزّ على جوانب مُحدَّدة من هذه الدراسات وبالقدر الذي يسمح لنا به هذا المؤلَّف مبتدئين ذلك بدراسة مبسطة للارتباط الخطي البسيط، وبعد ذلك الانتقال إلى مسائل متعلّقة بالانحدار، وأخيراً نختم هذا الفصل بعرض موجز عن الارتباط المتعدّد والجزئي. لكن قبل البدء في موضوع الارتباط الخطي البسيط سنقدّم بعض المفاهيم العامّة والمتعلّقة بالارتباط حيث يمكن لمن يودّ الاطلاع على المزيد الرجوع إلى المصادر المتخصِّصة في الارتباط وتحليل الانحدار والمذكور بعضها في فهرس المراجع في نهاية هذا الكتاب مثل [22] و[33] و[49] و[63].

لقد تحدثنا في الفصلين الأول والثاني عن طرائق إحصائية مُستخدمة في وصف ودراسة متغيِّر واحد، حيث لاحظنا أنَّه يُمكن لتلك الطرائق أن تُعطي تصوراً مقبولاً حول المتغيِّر الذي هو قيد الدراسة، ولكنَّها لا تقدِّم لنا تصوراً واضحاً ودقيقاً حول التّأثير المتبادل بين متغيِّرين أو أكثر. بمعنى أنَّه لو طُرح علينا السؤال الآتي:

**هل التّأثير المتبادل بين ظاهرتين ممثَّلتين بمتغيِّرين $X$ و$Y$، قوي أم ضعيف، طردي أم عكسي؟**

للإجابة على هذا السؤال لا بدّ لنا من منهج علمي دقيق وواضح بعيداً عن التخمين الحدسّي، ويقدّم لنا مقداراً عددياً يُعبِّر عن طبيعة التّأثير المتبادل بين المتغيِّرات.

إنَّ الارتباط (**وفقاً لعلم الإحصاء**) هو أحد المناهج العلميّة الذي يتولّى هذه المهمّة، ويهتمُّ بتعيين مقدار عددي يُعبِّر عن التّأثير المتبادل بين عدد من المتغيِّرات (**قـد يعود إلى وجود علاقة سببية بينها- مثل العلاقة بين أطوال الأشخاص وأوزانهم-**)، وهنا يُمكننا أنَّ نميِّز بين نماذج مختلفة من الارتباط كما يلي:

**أولاً: من حيث التّأثير والتّأثُّر:**

في هذا الإطار يوجد عدَّة أنواع من الارتباط بحسب العلاقة التي تصفه، ومن هذه العلاقات ما يلي:

**١- العلاقة الوحيدة الاتجاه Unidirectional Relation:**

هذا النوع من الارتباط مبني على أنَّ التغيُّر في قيمتغيِّر يؤدي إلى تغيّر في قيم متغيِّر آخر مرتبط به، وهذا التّأثير غير عكوس. أي إن التغيُّر في المتغيِّر الأخير لا يؤثِّر في تغيُّر المتغيِّر الأول، ولذلك يُدعى هذا النوع من الارتباط باسم **الارتباط السببي** Causal Correlation، ومن الأمثلة على ذلك:

أ- زيادة نسبة الأسمدة وفق معايير سليمة يؤدي إلى زيادةالإنتاج، إلاّ أنَّ العكس ليس صحيحاً بالضرورة.

ب- كلما ازداد ارتفاعنا عن مستوى سطح البحر تنخفض درجة حرارة الهواء عموما، إلاّ أنَّ العكس ليس صحيحاً بالضرورة.

**٢ – علاقة السلسلة السببية Causal Chain Relation:**

إنَّ هذا النوع من الارتباط يشابه النوع السّابق من الارتباط، ولكنَّه يكون في سلسلة من نماذج الارتباط الوحيد الاتجاه (السببي)، فعلى سبيل المثال: هطول الأمطار يؤدي إلى تسرب مياه الأمطار إلى التربة، وهذا يزيد المحتوى المائي للتربة الذي بدوره يتسرب إلى مخازن المياه الجوفية، ونحصل نتيجة لذلك على زيادة في مخزون المياه الجوفية وبهذا تشكَّل لدينا سلسلة من العلاقات وحيدة الاتجاه، ولذلك يُمكن إدراج هذا النوع من الارتباط تحت اسم **الارتباط السببي** أيضاً.

**٣ – العلاقة التبادلية Reciprocal Relation:**

إنَّ هذا النوع من الارتباط مبنيّ على التّأثير المتبادل بين المتغيِّرات أو على ردّ الفعل، فكل تَغيُّر في متغيِّر يؤدي إلى تَغيُّر في المتغيِّر الآخر، فعلى سبيل المثال، لو كان لدينا دارة كهربية مكوَّنة من مقاومة ومَنبع للتغذية (كما في الشكل التابع للمثال / ٥/ من /٢ , ٤ , ١ , ١/ في الفصل الأول) فعندئذ تغيُّر الجهد الكهربائي المطبَّق على طرفي الدارة سيؤدي إلى تغيُّر شدَّة التيّار المار في المقاومة، والعكس صحيح أيضاً، فإنَّ تغيّر شدَّة التيّار المار في المقاومة سيرافقه تغيُّراً في الجهد الكهربائي المُطبَّق على طرفي هذه المقاومة. لذلك أطلق على هذا النوع من الارتباط اسم **الارتباط التبادلي** Reciprocal Correlation.

**٤ – العلاقة الوهمية Spurious Relation**

في هذا النوع من الارتباط يبدو لنا ظاهريًّا نوع من التّأثير لمتغيِّر على متغيِّر آخر، ولكنَّ هذا التّأثير غير حقيقيّ، فعلى سبيل المثال: لو أخذنا تغيُّر مستوى الوعي لدى طفل ومقارنة ذلك مع تغيُّر طول يد هذا الطفل، فنجد أنَّ هناك علاقة طردية بينهما ولكنَّها ليست حقيقية، فلا تأثير لطول يد الطفل على وعيه ولا لوعيه على طول يده. لذلك أطلق على هذا النوع من الارتباط اسم **الارتباط الوهميّ** Spurious Correlation.

**ثانياً: من حيث عدد الظواهر**

يمكننا هنا أن نصنِّف نوعين رئيسيين هما:

**١ – الارتباط البسيط Simple Correlation**

إنَّ **الارتباط البسيط** هو الارتباط الذي يهتمّ بالتّأثير المتبادل بين متغيِّرين فقط، ومثال على ذلك تعيين مقدار التّأثير المتبادل بين طول شخص ما ووزنه.

**٢ – الارتباط المتعدِّد Multiple Correlation**

إنَّ الارتباط المتعدِّد هو الارتباط الذي يهتمّ بالتّأثير المتبادل بين عدَّة متغيِّرات (أكثر من متغيِّرين)، كأن نبحث في التّأثير المتبادل بين ضغط الدم، ودرجة الحرارة، وعدد ضربات القلب، وكمية الأكسجين في الدم لمرضى السكر، فنكون في هذه الحالة أمام ما يُعرف باسم **الارتباط المتعدد**، وفي هذا النوع من الارتباط قد يلجأ الباحث إلى دراسة التّأثير المتبادل بين متغيِّرين فقط مع تجاهل (أو تثبيت) التّأثير المتبادل بين بقية المتغيِّرات، وفي هذه الحالة نكون أمام ما يُعرف باسم **الارتباط الجزئي** Partial Correlation.

الآن لنفترض أنَّه لدينا $(x_1,y_1)$ ، $(x_2,y_2)$ ،...و $(x_n,y_n)$ بيانات مقدَّمة على شكل ثنائيات مرتبة (ندعوها بيانات متزاوجة). إنَّ طبيعة هذه البيانات لا تهمنا بشيء، فقد تكون طبيعة البيانات $x_1$ ، $x_2$ ،... و $x_n$ مختلفةً عن طبيعة البيانات $y_1$ ، $y_2$ ،... و $y_n$ ، ولكن المهمّ هنا أن يكون البيان $y_i$ هو البيان المقابل لـ $x_i$ من أجل كل القيم الممكنة لـ $i$ ، فعلى سبيل المثال لو أخذنا البيانات الآتية التي تمثِّل نتائج الاختبار التحريري (مقدَّرةً بالدرجات) والمقابلة (مقدّمةً بالتقديرات: ممتاز – جيد جداً – جيد – وسط – ضعيف) في مقرَّر الإحصاء لعشرة من الطلاب كما في الجدول الآتي:

الجدول (٣,١)

| الطالب | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| الامتحان التحريري $X$ | 8 | 8 | 5 | 9 | 7 | 7 | 9 | 10 |
| امتحان المقابلة $Y$ | جيد | جيد جداً | وسط | جيد | جيد | وسط | جيد | جيد جداً |

فنلاحظ أنَّ البيانات $x_1$، $x_2$،... و $x_n$ هي بيانات كميّة في حين أنَّ البيانات $y_1$ ، $y_2$ ،... و $y_n$ هي بيانات نوعيّة. أكثر من ذلك قد تكون هذه البيانات ناتجة عن تطبيق دراسة متعلّقة بظاهرتين $A$ و$B$ على عناصر عيّنة أو مجتمع كما في المثال السّابق حيث لدينا الظاهرة $A$ هي نتائج الاختبار التحريري والظاهرة $B$ هي نتائج اختبار المقابلة. لا بل أكثر من ذلك فقد تكون البيانات $x_1$، $x_2$،... و $x_n$ هي لحظات زمنية متتالية في حين أنَّ $y_1$ ، $y_2$ ،... و $y_n$ هي قيم لظاهرة $B$ متعلّقة بتطور هذه اللحظات الزمنية المتتالية على الترتيب، فعلى سبيل المثال لو أخذنا البيانات المقدَّمة في الجدول الآتي (تمثّل إنتاج القمح مقدّراً بالـ 100,000 طن في بلد ما من العام 2002 وحتى 2010):

الجدول (٣,٢)

| السّنة | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 | 2009 | 2010 |
|---|---|---|---|---|---|---|---|---|---|
| الكميّة | 8 | 12 | 9 | 15 | 11 | 9 | 7 | 9 | 13 |

ومن الممكن أن تكون البيانات مولّدة بطرائق أخرى مختلفة عن الحالات المقدَّمة سابقاً أيضاً، والمهمّ هنا (وكما ذكرنا سابقاً) أن تكون البيانات المُعطاة متزاوجة.

الآن، ولدراسة الارتباط بين متغيّرين $X$ و$Y$، سنميّز بين الحالتين الآتيتين:

١ – الحالة التي تكون فيها العلاقة التي تربط بين المتغيّرين $X$ و$Y$ هي علاقة خطيّة فقط، أي من الشكل:

$$Y = f(X) = a + bX$$

علماً أنَّ $a$ و$b$ قيمتان ثابتتان يُطلب تعيينهما، فعندئذ يتحدث المرء عن **الارتباط الخطي البسيط**.

٢ – الحالة التي تكون فيها العلاقة التي تربط بين المتغيّرين $X$ و$Y$ هي علاقة غير خطيّة، كأن تكون العلاقة بينهما (على سبيل المثال لا الحصر) من أحد الأشكال الآتية:

$$Y = f(X) = a \sin bX \qquad \& \qquad Y = f(X) = a \cdot b^X$$

$$Y = f(X) = a_o + a_1 X + a_2 X^2 + \dots + a_m X^m \quad ; m \geq 2$$

علماً أنَّ $a$، $b$ و$c$، وكذلك $a_o$ ، $a_1$ ، $a_2$ ،... و $a_n$ قيم ثابتة يُطلب تعيينها، فعندئذ يتحدث المرء عن **الارتباط غير الخطي** Non Linear Correlation.

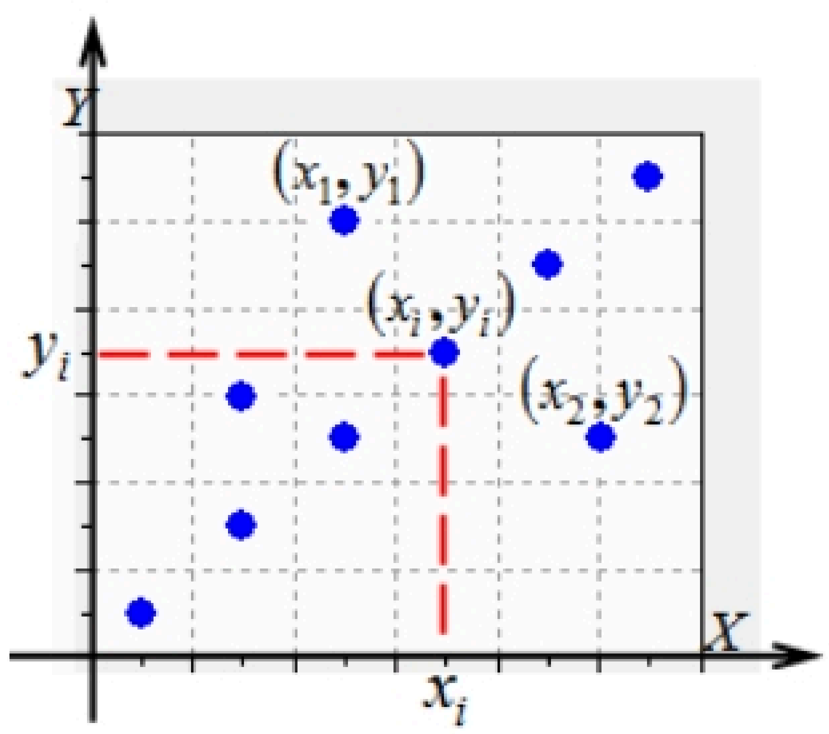
### (٣,١,١,١) ملاحظة

تجدر الملاحظة هنا إلى أنَّ درجة الارتباط التي نحصل عليها تعتمد على نوع العلاقة المفترضة ومدى صحة هذا الافتراض، وهذه حقيقة يجب مراعاتها عند تفسير النتائج التي نحصل عليها، فلو افترضنا أنَّ العلاقة بين متغيّرين هي علاقة خطيّة، وتبيّن لنا أنَّ درجة الارتباط (أي قوة الارتباط، أهي: قوية، أم متوسطة، أم ضعيفة) منخفضة جداً فإنَّ ذلك لا يعني أكثر من احتمال عدم وجود علاقة خطيّة بين المتغيّرين، ولا يعني عدم وجود ارتباط بينهما، إذ إنَّه من المُمكن أن يكون الارتباط بينهما له شكل علاقة غير خطيّة. كما أنَّ وجود درجة ارتباط قوية بينهما قد لا يعني وجود علاقة أصلاً بين الظاهرتين مثلما هو الحال لدى الارتباط الوهمي.

الآن وقبل تقديم بعض المعايير العدديّة التي تقيس لنا درجة الارتباط بين متغيّرين $X$ و $Y$ سنقوم بشرح لمفهوم **لوحة الانتشار** التي تُعدّ من الوسائط المهمّة في هذه الدراسات. إنَّ الغاية من هذه اللوحة أخذ انطباع أولي عن طبيعة التّأثير المتبادل بين المتغيّرين التفسيريّ $X$ والاستجابة $Y$.

### (٣ ,١ ,٢) لوحة الانتشار Scatter Plot

لتكن لدينا $(x_1,y_1)$ ، $(x_2,y_2)$ ،...و $(x_n,y_n)$ بيانات مُعطاة، فعندئذ لوحة الانتشار لهذه البيانات هي التمثيل النقطي للبيانات (أي تمثيل البيانات بنقاط) على المستوي الإحداثي **XoY** ، ويُؤخذ عادةً في الإحداثيات المتعامدة، أي إنَّ $x_1$ ، $x_2$ ،... و $x_n$ وكذلك $y_1$ ، $y_2$ ،... و $y_n$ هي مساقط النقاط الممثّلة لهذه البيانات على المحور **oX** و **oY** على الترتيب (انظر الشكل الجانبي)، ولإعطاء المزيد من التوضيح سنقدِّم الأمثلة الآتية.



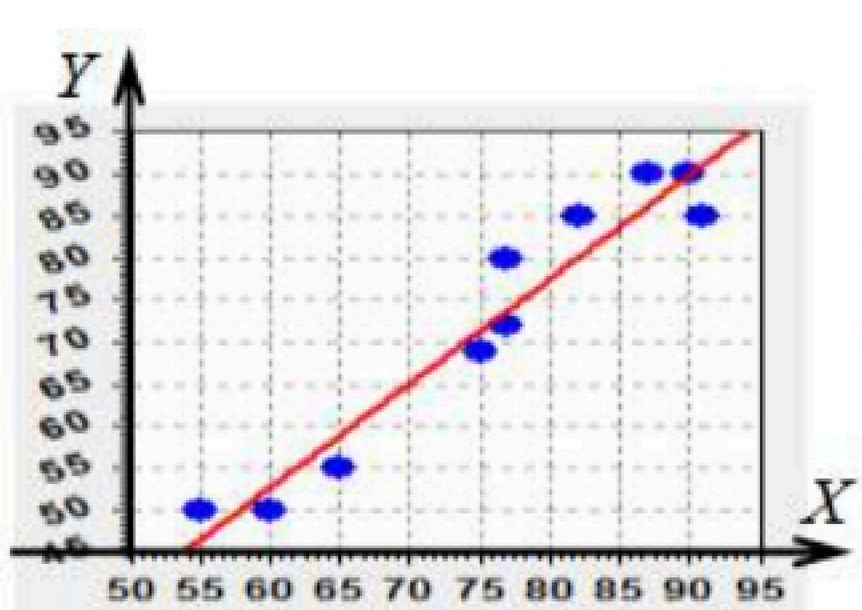

#### (٣ ,١ ,٢ ,١) أمثلة:

١ – لنفترض أنّنا نرغب في الكشف عن درجة الارتباط بين الطول والوزن لعشرة أشخاص بالغين، ومن أجل هذه الغاية قمنا بقياس الطول والوزن لعشرة أشخاص بالغين تمَّ اختيارهم عشوائيًّا، فحصلنا على البيانات $(x_1,y_1)$ ، $(x_2,y_2)$ ،...و $(x_{10},y_{10})$ المقدَّمة من خلال الجدول الآتي:

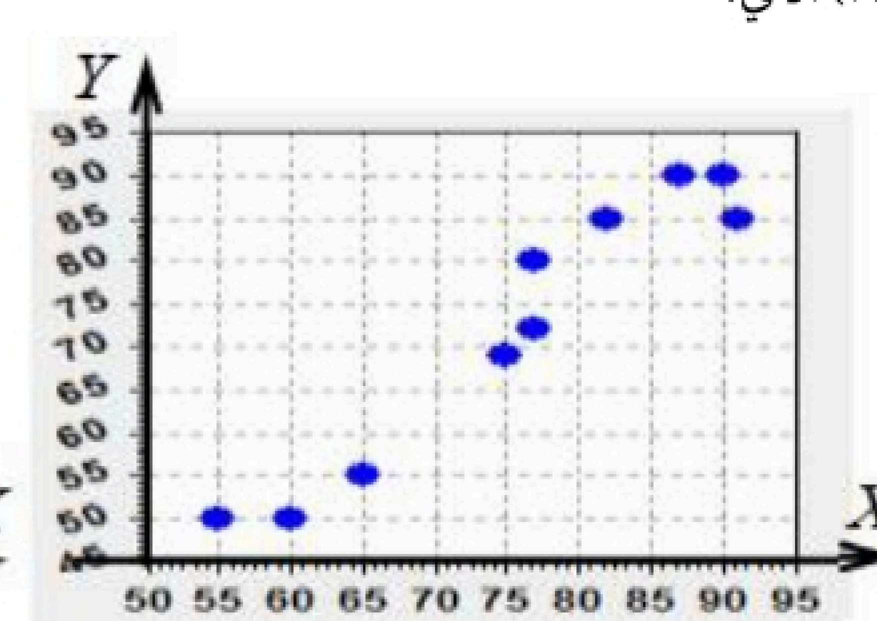
الجدول (٣ ,٣)

| $i$ | الوزن $X$ | الطول $Y$ | $i$ | الوزن $X$ | الطول $Y$ |
|---|---|---|---|---|---|
| 1 | 75 | 69 | 6 | 60 | 50 |
| 2 | 82 | 85 | 7 | 55 | 50 |
| 3 | 65 | 55 | 8 | 87 | 90 |
| 4 | 90 | 90 | 9 | 91 | 85 |
| 5 | 77 | 80 | 10 | 77 | 72 |

علماً أنَّ الوزن $X$ مقدَّر بالكيلو غرام والطول $Y$ مقدَّر بالبو صة Inch. عندئذ نجد أنَّ لوحة الانتشار لهذه البيانات لها العرض المو ضَّح بالشكل (١ ,٣. أ) الآتي.



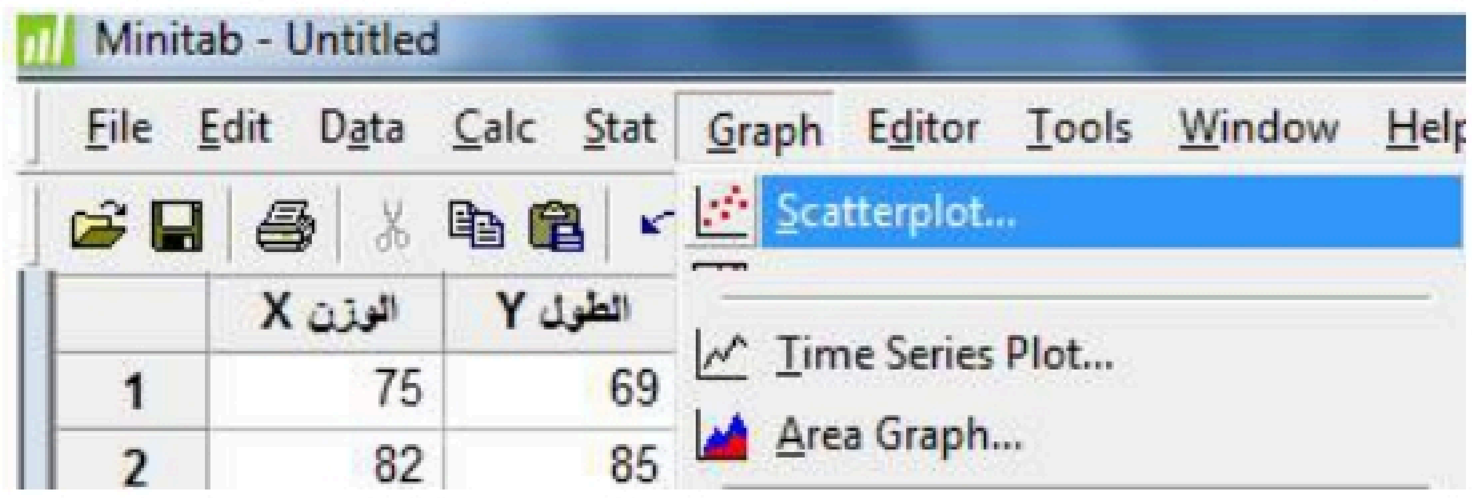

الشكل (١ ,٣. ب)



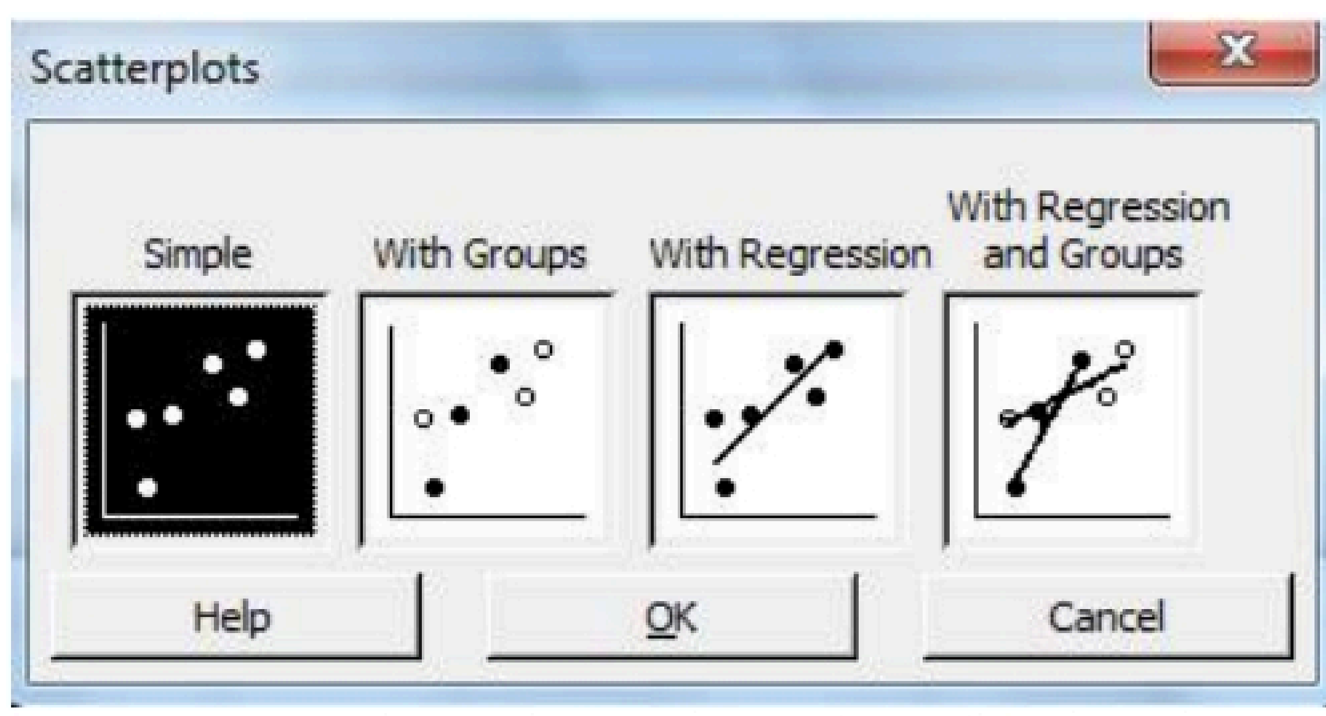

الشكل (١ ,٣. أ)

حيث نلاحظ أنَّ هذه النقاط تُظهر مَنْحى صاعداً، بمعنى أنّه كلما ازدادت قيمة مُركّبة أدى ذلك إلى زيادة قيمة المُركّبة الأخرى، وهذا السلوك في الارتباط يُدعى **الارتباط الإيجابي** (أو الطردي) للبيانات. أمّا بخصوص قوة التّأثير المتبادل (أي قوة الارتباط وسوف نوضّح لاحقاً معنى قوة الارتباط) بين الطول والوزن، فإنَّ الشكل السّابق يُظهر لنا درجةً عاليةً من التّأثير المتبادل بين الطول والوزن وذلك بسبب وقوع معظم البيانات بالقرب من مستقيم مُمثِّل لمنحى هذه البيانات (سنأتي على شرح كيفية تعيين هذا المستقيم لاحقاً في تحليل الانحدار الخطي من هذا الفصل)، انظر الشكل السّابق (١ ,٣ .ب).

يمكننا الحصول على لوحة الانتشار (أو لوحة الانتشار مع الخط المستقيم المار بين البيانات) باستخدام برنامج Minitab من خلال الفقرة Graph ثمَّ Scatterplot وبعد ذلك اختيار المربع Simple أو المربع With Regression، وأخيراً OK، والصور الآتية توضِّح ذلك:




ومن ثمَّ



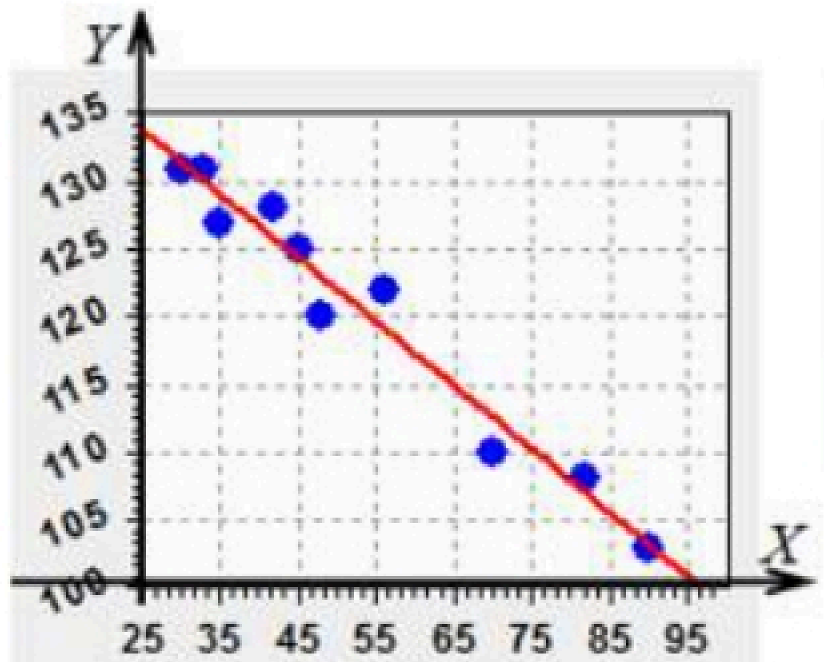

فنحصل على الشكلين السابقين (١ , ٣ .أ) و (١ , ٣ .ب).

٢ - أخضعت مجموعة مكوَّنة من 10 رجال بدينين من ذوي وزن 140 كغ لدورة رياضية ونظام غذائي موحّد من أجل تخفيض أوزانهم بحيث يمارس كل منهم فترة زمنية ما يومياً حسب استطاعته ولمدّة 45 يوماً، وبعد ذلك قيس الوزن لكلٍّ منهم، ومن ثمَّ دوِّن إلى جانب عدد السّاعات التي مارسها، فكانت النتائج كما في الجدول الآتي:

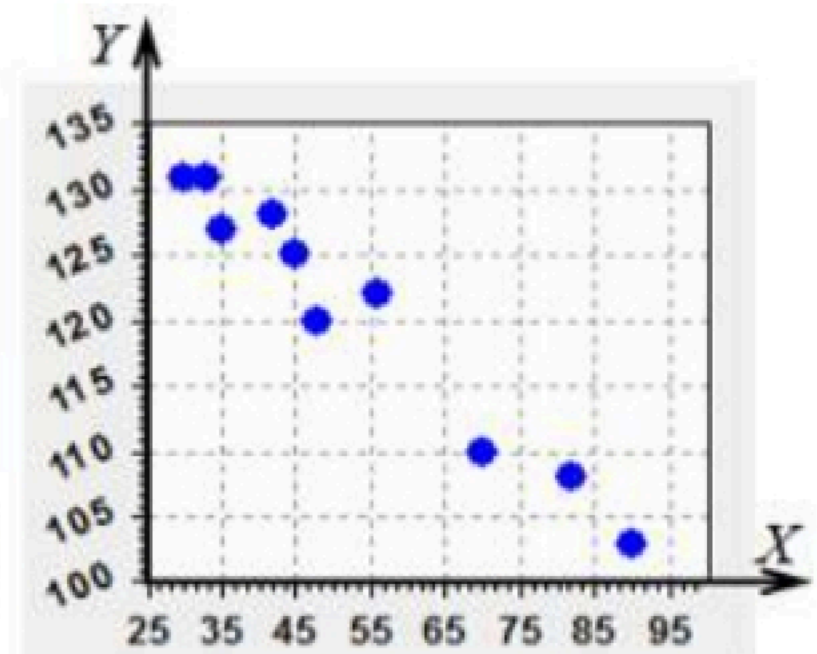
الجدول (٤ , ٣)

| الرَّجُل | $X$ عدد الساعات التي مارسها خلال 45 يوماً | $Y$ الوزن عند نهاية الدورة |
|---|---|---|
| 1 | 30 | 131 |
| 2 | 45 | 125 |
| 3 | 56 | 122 |
| 4 | 42 | 128 |
| 5 | 90 | 103 |
| 6 | 33 | 131 |
| 7 | 48 | 120 |
| 8 | 35 | 127 |
| 9 | 70 | 110 |
| 10 | 82 | 108 |

فنجد أنَّ لوحة الانتشار للبيانات المُعطاة لها العرض الموضَّح بالشكل (٢ , ٣ .أ) الآتي:




الشكل (٢, ٣ .ب)




الشكل (٢, ٣ .أ)

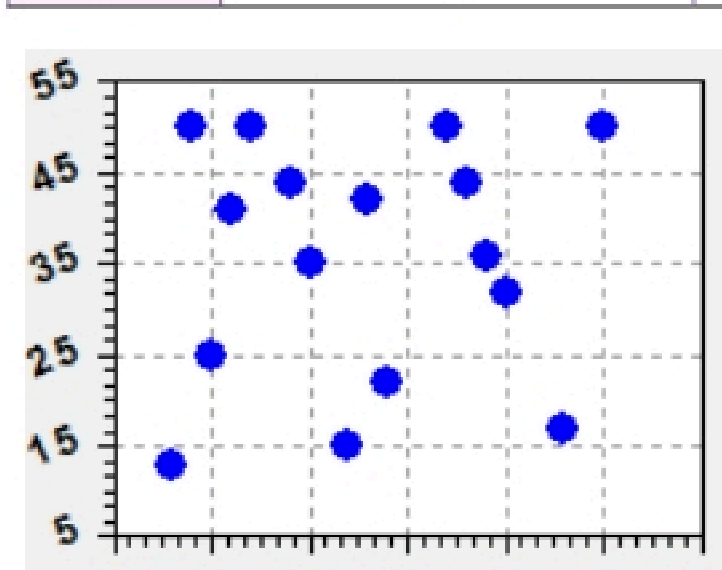
وكذلك نلاحظ أنَّ النقاط الممثِّلة للبيانات تُظهر مَنْحى هابطًا بحيث إنه كلّما ازدادت قيمة مُركِّبة أدى ذلك إلى نقصان في قيمة المُركِّبة الأخرى، وهذا السلوك في الارتباط يُدعى **الارتباط السلبي** (أو العكسي) للبيانات، وأمّا بخصوص درجة الارتباط فإنَّ الشكل يظهر لنا درجةً عاليةً من التّأثير المتبادل بين عدد ساعات الرياضة ونقصان الوزن، وذلك بسبب وقوع معظم البيانات بالقرب من المستقيم الممثِّل لمنحى هذه البيانات (انظر الشكل (٣, ٢ .ب)).

٣- تقدَّم ستة عشر طالباً للاختبار النهائي في مقرَّري الرياضيات واللغة العربية، فكانت لهم النتائج كما في الجدول الآتي (مقدَّرةً من 50 درجة).

الجدول (٥, ٣)

| $Y$ درجة اختبار اللغة العربية | $X$ درجة اختبار الرياضيات | الطالب | $Y$ درجة اختبار اللغة العربية | $X$ درجة اختبار الرياضيات | الطالب |
|---|---|---|---|---|---|
| 13 | 28 | 9 | 32 | 45 | 1 |
| 35 | 35 | 10 | 17 | 48 | 2 |
| 22 | 39 | 11 | 25 | 30 | 3 |
| 50 | 42 | 12 | 50 | 50 | 4 |
| 50 | 32 | 13 | 44 | 43 | 5 |
| 42 | 38 | 14 | 15 | 37 | 6 |
| 41 | 31 | 15 | 44 | 34 | 7 |
| 50 | 29 | 16 | 36 | 44 | 8 |

عندئذ نجد أنَّ لوحة الانتشار لهذه البيانات لها العرض الموضَّح بالشكل (٣, ٣)، وكذلك نلاحظ أنَّ النقاط الممثِّلة للبيانات المُعطاة تُظهر توزعًا على شكل غيمة يصعب تحديد اتجاهها (المَنْحى غير واضح تماماً)، بمعنى أنَّه من خلال النظر بالعين المجرّدة يصعب تحديد طبيعة التّأثير المتبادل بين المتغيِّرين $X$ و$Y$ من حيث الإيجابية أو السلبية.




الشكل (٣, ٣)

إذاً، وكما لاحظنا، فإنَّنا بحاجة لمقياس عددي يوضّح لنا طبيعة الارتباط أهو إيجابي أم سلبي؟ أم هو قوي أم ضعيف؟ إنَّ المقاييس العددية التي تقوم بهذه المهمَّة تُدعى **مُعاملات الارتباط** Correlation Coefficients، وسنرمز له بـ $\mathbf{r}$ من أجل العينات وبـ$\rho$ من أجل المجتمعات، وقد تُرفق هذه الرموز بدليل لتمييز نوع المقياس المستخدم.

**(٢ , ٢ , ١ , ٣) ملاحظات**

قبل القيام بتقديم بعض المفاهيم المتعلّقة بالارتباط والانحدار، فإنّنا سنتّفق على الآتي تجنباً للتقديم المطوّل، ومن أجل توضيح المفاهيم التي سنقدّمها:

١- عندما نذكر أنّه لدينا $(x_1, y_1)$، $(x_2, y_2)$، ...و $(x_N, y_N)$ بيانات مجتمع خاضع لمتغيّرين $X$ و $Y$، فإنّنا نقصد بذلك:

أ- إمّا أن تكون عناصر مجتمع إحصائي منته حجمه $N$ أخضعت لدراسة متّجه $(X, Y)$ (مكوّناته المتغيّرين $X$ و $Y$) وحصلنا نتيجةً لذلك على هذه البيانات المتزاوجة، ومثال على ذلك تنفيذ مسح شامل لمجتمع مرضى السكري في مدينة ما بحيث يُحدَّد لكلّ مريض نسبة السكر والشحوم الثلاثية في دمه، فعندئذ نحصل على بيانات متزاوجة من النموذج الموصوف آنفًا.

ب- أو أن تكون البيانات $x_1$، $x_2$، ... و $x_N$ هي قيم ناتجة عن إخضاع عناصر مجتمع إحصائي منته حجمه $N$ لدراسة المتغيّر $X$، في حين أنّ البيانات $y_1$، $y_2$، ... و $y_N$ هي قيم ناتجة عن إخضاع عناصر مجتمع إحصائي آخر منته حجمه $N$ لدراسة المتغيّر $Y$، وبحيث تُكوِّن البيانات $x_1$، $x_2$، ... و $x_N$ و $y_1$، $y_2$، ... و $y_N$ البيانات المتزاوجة المذكورة آنفًا، وكمثال على ذلك لو رصدنا قيم البورصة لإحدى الشركات منذ اشتراكها في البوصة وحتى وقت بدء هذه الدراسة وعلى فترات زمنية منفصلة ومتتالية (ليس بالضرورة أن تكون الفترات الزمنية متساوية)، فعندئذ نحصل على بيانات متزاوجة من النموذج الموصوف آنفًا. حيث تلعب البيانات $x_1$، $x_2$، ... و $x_N$ دور اللحظات الزمنية التي رصدت عندها قيم البورصة $y_1$، $y_2$، ... و $y_N$ .

٢- عندما نذكر أنّه لدينا $(x_1, y_1)$، $(x_2, y_2)$، ...و $(x_n, y_n)$ بيانات عيّنة خاضعة لمتغيّرين $X$ و $Y$، فإنّنا نقصد بذلك:

أ- إمّا أن تكون عناصر عيّنة حجمها $n$ مسحوبة من مجتمع إحصائي أخضعت لدراسة هذين المتغيّرين $X$ و $Y$ وحصلنا نتيجةً لذلك على هذه البيانات المتزاوجة المذكورة آنفًا.

ب- أو أن تكون البيانات $x_1$، $x_2$، ... و $x_n$ هي قيم ناتجة عن إخضاع عناصر عيّنة حجمها $n$ مسحوبة من مجتمع إحصائي لدراسة المتغيّر $X$، وكذلك البيانات $y_1$، $y_2$، ... و $y_n$ هي قيم ناتجة عن إخضاع عناصر عيّنة حجمها $n$ مسحوبة من مجتمع إحصائي آخر لدراسة المتغيّر $Y$، وبحيث تُكوِّن البيانات $x_1$، $x_2$، ... و $x_n$ وكذلك $y_1$، $y_2$، ... و $y_n$ هي البيانات المتزاوجة المذكورة آنفًا.

فيما يلي نقدّم أحد التعاريف المهمّة في إطار الارتباط وتحليل الانحدار، ويُعرف باسم **التغاير** Covariance (أو تمام التباين).

**(٣ , ١ , ٣) تعريف (التغاير لمجموعتي بيانات Covariance of two Data Sets)**

من أجل تعريف التغاير لمجموعتين من البيانات سنناقش الحالتين الآتيتين:

**أولاً: من أجل البيانات الخام:**

أ- لنفترض أنّ $(x_1, y_1)$، $(x_2, y_2)$، ...و $(x_N, y_N)$ بيانات مجتمع خاضع لمتغيّرين $X$ و $Y$، فعندئذ يُعرَّف **التغاير** بين بيانات المتغيّرين $X$ و $Y$ (سنرمز له بـ $\sigma_{X,Y}$) بالعلاقة الآتية:

$$\sigma_{X,Y} := \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_X)\cdot(y_i - \mu_Y) \qquad [3,1\text{-a}]$$

علماً أنّ $\mu_X$ و $\mu_Y$ هو المتوسط لبيانات $X$ و $Y$ على الترتيب.

ب- لنفترض أنّ $(x_1, y_1)$، $(x_2, y_2)$، ...و $(x_n, y_n)$ بيانات عيّنة خاضعة لمتغيّرين $X$ و $Y$، فعندئذ يُعرَّف **التغاير** بين بيانات $X$ و $Y$ (سنرمز له بـ $S_{X,Y}$) بوساطة العلاقة الآتية:

$$S_{X,Y} := \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})\cdot(y_i - \overline{y}) \qquad [3,1\text{-b}]$$

علماً أنّ $\overline{x}$ و $\overline{y}$ هو المتوسط لبيانات $X$ و $Y$ على الترتيب.

**ثانياً: إذا كانت البيانات المُعطاة مُجمَّعة في جدول تكراري مشترك:**

إذا كان لدينا مجموعتي بيانات إحداها ناتجة عن متغيِّر $X$ والأخرى ناتجة عن متغيِّر $Y$، ومُعطاة من خلال جدول تكراري واحد ذي فئات تَظهر فيه التكرارات للبيانات المشتركة بين مجموعتي البيانات، ويُدعى هذا الجدول بـ **الجدول التكراري المشترك** Joint Frequency Table لمجموعتي البيانات.

الآن لنفترض أنَّه لدينا مجموعتين من البيانات إحداها ناتجة عن متغيِّر $X$ والأخرى ناتجة عن متغيِّر $Y$، وصُبَّت هذه البيانات في جدول تكراري مشترك كما يلي:

الجدول (٦ , ٣)

| | | فئات $X$ | | | | Total |
|---|---|---|---|---|---|---|
| | | $\alpha_1 - \alpha_2$ | $\alpha_2 - \alpha_3$ | ......... | $\alpha_k - \alpha_{k+1}$ | |
| فئات $Y$ | $\beta_1 - \beta_2$ | $f_{11}$ | $f_{21}$ | ......... | $f_{k1}$ | $\sum_{i=1}^{k} f_{i1}$ |
| | $\beta_2 - \beta_3$ | $f_{12}$ | $f_{22}$ | ......... | $f_{k2}$ | $\sum_{i=1}^{k} f_{i2}$ |
| | $\vdots \quad \vdots$ | ......... | ......... | ......... | ......... | ......... |
| | $\beta_\ell - \beta_{\ell+1}$ | $f_{1\ell}$ | $f_{2\ell}$ | ......... | $f_{k\ell}$ | $\sum_{i=1}^{k} f_{i\ell}$ |
| Total | | $\sum_{j=1}^{\ell} f_{1j}$ | $\sum_{j=1}^{\ell} f_{2j}$ | ......... | $\sum_{j=1}^{\ell} f_{kj}$ | $\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}$ |

علماً أنَّ $f_{ij}$ هو التكرار المشترك بين الفئة $[\alpha_i , \alpha_{i+1})$ و $[\beta_j , \beta_{j+1})$، فعندئذ:

١- إذا كانت هذه البيانات لمجتمعين إحصائيين منتهيين حجم كل منها $N$، فعندئذ يُحسب التغاير بين بيانات $X$ و$Y$ والممثَّلة بالجدول السّابق (٣ , ٤) كما يلي:

$$\sigma_{X,Y} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}\,(x_i - \mu_X)(y_j - \mu_Y)}{\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}} \qquad [3,2\text{-a}]$$

علماً أنَّ $x_i$ هو مركز الفئة $\alpha_i \to \alpha_{i+1}$ من أجل كل $i \in \mathbb{N}_k$، وبالمثل $y_i$ هو مركز الفئة $\beta_j \to \beta_{j+1}$ من أجل كل $j \in \mathbb{N}_\ell$، وأخيراً $\mu_X$ و $\mu_Y$ هو المتوسط لبيانات $X$ و$Y$ على الترتيب، ويُحسبان كما يلي:

$$\mu_X = \frac{\sum_{i=1}^{k}\left(x_i \cdot \sum_{j=1}^{\ell} f_{ij}\right)}{\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}} \qquad [3,2\text{-b}]$$

$$\mu_Y = \frac{\sum_{j=1}^{\ell}\left(y_j \cdot \sum_{i=1}^{k} f_{ij}\right)}{\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}} \qquad [3,2\text{-c}]$$

٢- إذا كانت هذه البيانات لعينتين حجم كل منهما $n$، فعندئذ يُحسب التغاير بين بيانات $X$ و$Y$ والممثَّلة بالجدول السّابق (٣,٦) كما يلي:

$$S_{X,Y}=\frac{\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}\left(x_i-\overline{x}\right)\left(y_j-\overline{y}\right)}{\left(\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}\right)-1} \qquad [3,3\text{-a}]$$

علماً أنَّ $\overline{x}$ و $\overline{y}$ هو المتوسط لبيانات $X$ و $Y$ على الترتيب، ويُحسبان كما يلي:

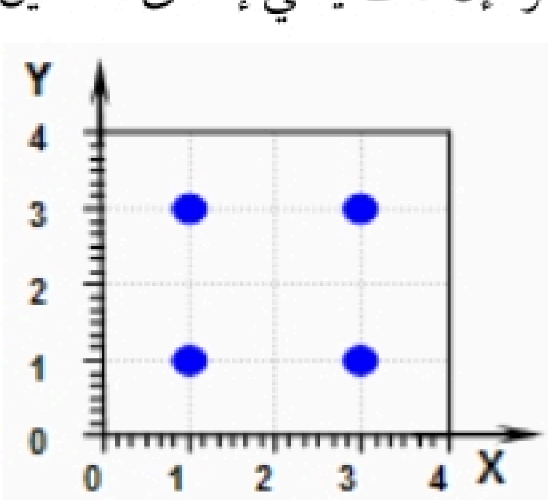
$$\overline{x}=\frac{\sum_{i=1}^{k}\left(x_i\cdot\sum_{j=1}^{\ell} f_{ij}\right)}{\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}} \qquad [3,3\text{-b}]$$

$$\overline{y}=\frac{\sum_{j=1}^{\ell}\left(y_j\cdot\sum_{i=1}^{k} f_{ij}\right)}{\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}} \qquad [3,3\text{-c}]$$

فلاحظ أنَّ التغاير المُعطى بالعلاقات [3,2-a] و [3,3-a] هو التغاير الذي قدِّم سابقاً من أجل البيانات الخام نفسه، ولكن عُدِّلت صيغته لكي توافق الجداول التكرارية المشتركة.

### (١,٣,١,٣) ملاحظات

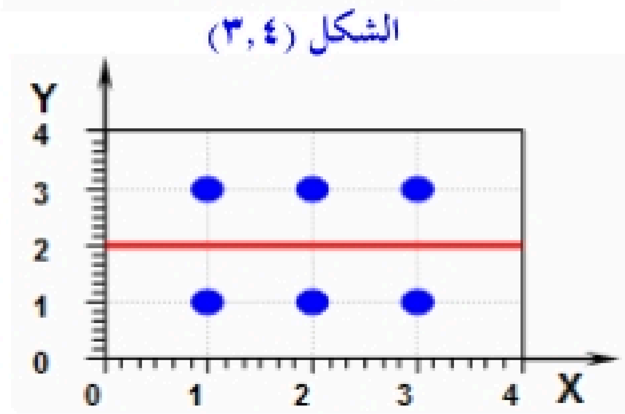
١- إذا وجدنا أنَّ قيمة الـتغاير موجبة تماماً، فإنَّ ذلك يعني أنَّ مَنْحى البيانات سيكون صاعداً، وأمّا إذا كانت قيمة التغاير سالبة تماماً فإنَّ ذلك يعني أنَّ مَنْحى البيانات سيكون هابطاً، وأخيراً إذا كانت قيمة التغاير تساوي الصفر فإنَّ ذلك يعني إحدى الحالتين الآتيتين:



الشكل (٣,٤)

أ ) يكون توزع البيانات في لوحة الانتشار ليس له مَنْحى واضح (أو على شكل غيمة عديمة الاتجاه)، فلو كانت (1,1)، (1,3)، (3,1) و(3,3) بيانات عيِّنة، فعندئذ نلاحظ أنّه لا يمكن تحديد مَنْحى لهذه البيانات، وبحساب التغاير بين قيم $X$ و$Y$ نجد:

$$S_{X,Y}:=\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\cdot\left(y_i-\overline{y}\right)=\frac{0}{3}=0$$



الشكل (٣,٥)

ب) يكون مَنْحى البيانات موازياً للمحور الأفقي (أو موازياً للمحور الرأسي)، فلو كانت (1,1)، (1,3)، (3,1)، (2,1)، (2,3) و(3,3) بيانات عيِّنة، فعندئذ نلاحظ أنّ مَنْحى هذه البيانات أفقي، وبحساب التغاير بين قيم $X$ و$Y$ نجد:

$$S_{X,Y}:=\frac{1}{n-1}\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\cdot\left(y_i-\overline{y}\right)=\frac{0}{5}=0$$

٢- إذا أخذنا الحالة الخاصّة $X=Y$، فعندئذ سينتج لدينا $S_{X,X}=S_X^2$ وهو تباين العيِّنة الناتجة عن تطبيق المتغيِّر $X$، وكذلك $\sigma_{X,X}=\sigma_X^2$ هو تباين المجتمع الناتج عن تطبيق المتغيِّر $X$ أيضاً.

٣- المصفوفة التي لها العرض الآتي:

$$\begin{pmatrix} S_{X,X} & S_{X,Y} \\ S_{Y,X} & S_{Y,Y} \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \sigma_{X,X} & \sigma_{X,Y} \\ \sigma_{Y,X} & \sigma_{Y,Y} \end{pmatrix}$$

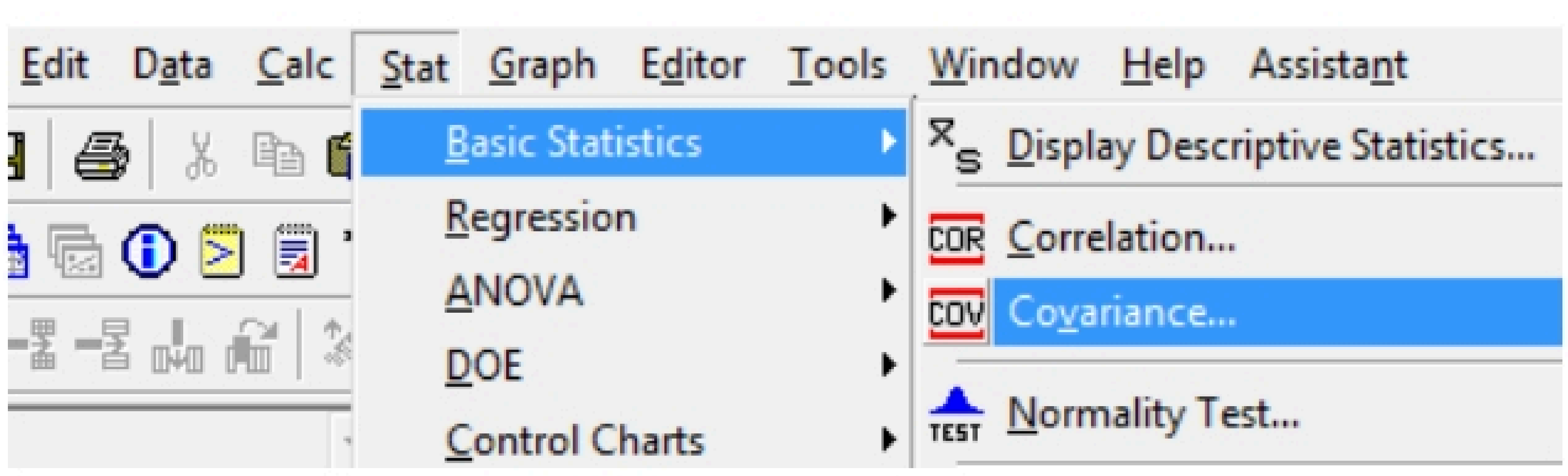
تُدعى **مصفوفة التغاير** Covariance Matrix لبيانات $X$ و$Y$، وهي مصفوفة متناظرة لأنَّ $S_{X,Y} = S_{Y,X}$ وكذلك $\sigma_{X,Y} = \sigma_{Y,X}$.

٤- عندما يُطلب من برنامج إحصائي (مثل Minitab أو SPSS) حساب التغاير للبيانات المقدّمة فإنَّه يتعامل معها على أنَّها بيانات عيِّنة فقط، فعلى سبيل التوضيح لو رجعنا إلى الأمثلة (٣-١-٢-١) فإنَّنا نجد ما يلي:

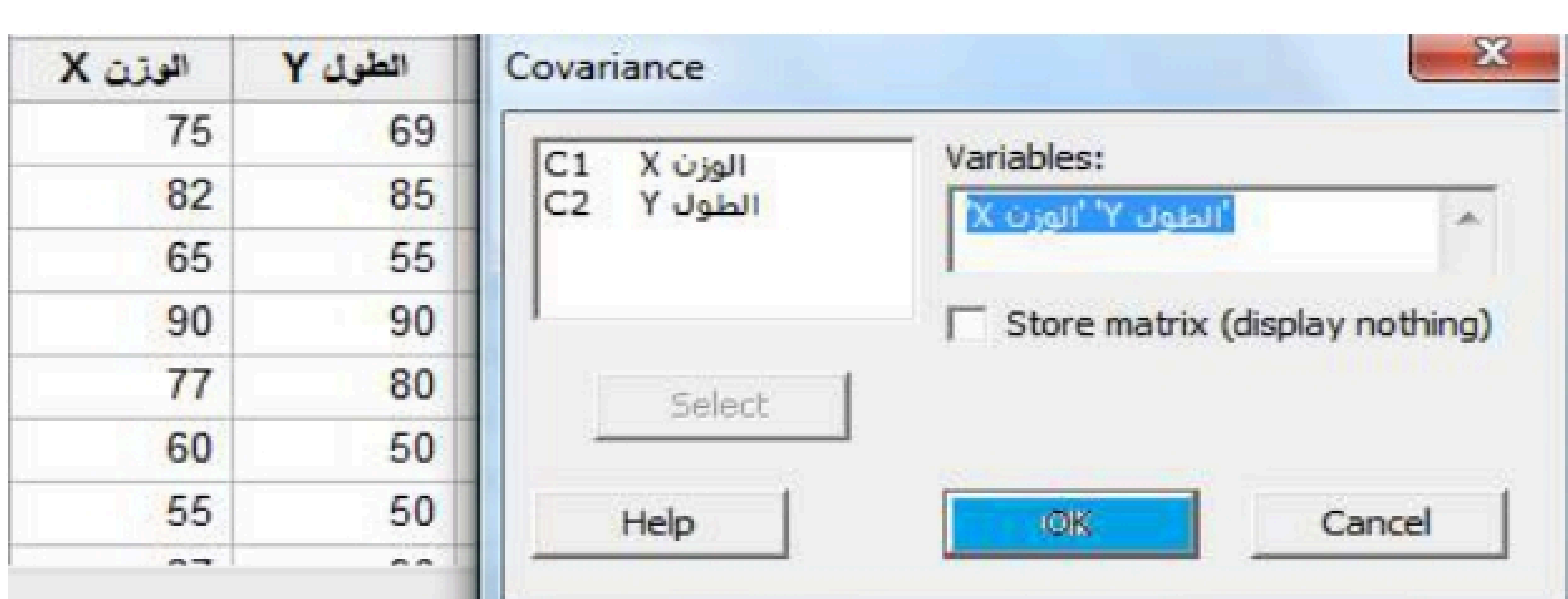
أ- من أجل بيانات المثال / ١/ من (١-٢-١-٣) لدينا $\bar{x} = 75.9$ وكذلك $\bar{y} = 72.6$، ومن ثمَّ تكون قيمة تغاير البيانات $X$ و$Y$ تساوي:

$$S_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})\cdot(y_i - \bar{y}) = \frac{1735.6}{9} = 192.844$$

وهكذا نجد $0 < S_{X,Y}$ ، وهذا يعني أنَّ مَنْحى البيانات يجب أن يكون صاعداً (وهذا موافق لما لاحظناه سابقاً على لوحة الانتشار في الشكل (٣,١))، ولدى استخدام برنامج Minitab لحساب قيم التغاير بين بيانات $X$ و$Y$ من خلال عرض مصفوفة التغاير لهما وذلك وفقاً للخطوات الموضَّحة في الأشكال الآتية:




ومن ثمَّ




فنحصل على التقرير الآتي لمصفوفة التغاير للبيانات المُعطاة:

**Covariances: X الوزن; Y الطول**

| | X الوزن | Y الطول | |
|---|---|---|---|
| X الوزن | 155.433 | | $S_X^2$ , $S_{X,Y}$ |
| Y الطول | 192.844 | 256.933 | $S_Y^2$ |

ب- من أجل بيانات المثال / ٢/ من (١-٢-١-٣) لدينا $\bar{x} = 53.1$ وكذلك $\bar{y} = 120.5$ ، ومن ثمَّ تكون قيمة التغاير بين البيانات $X$ و$Y$ تساوي:

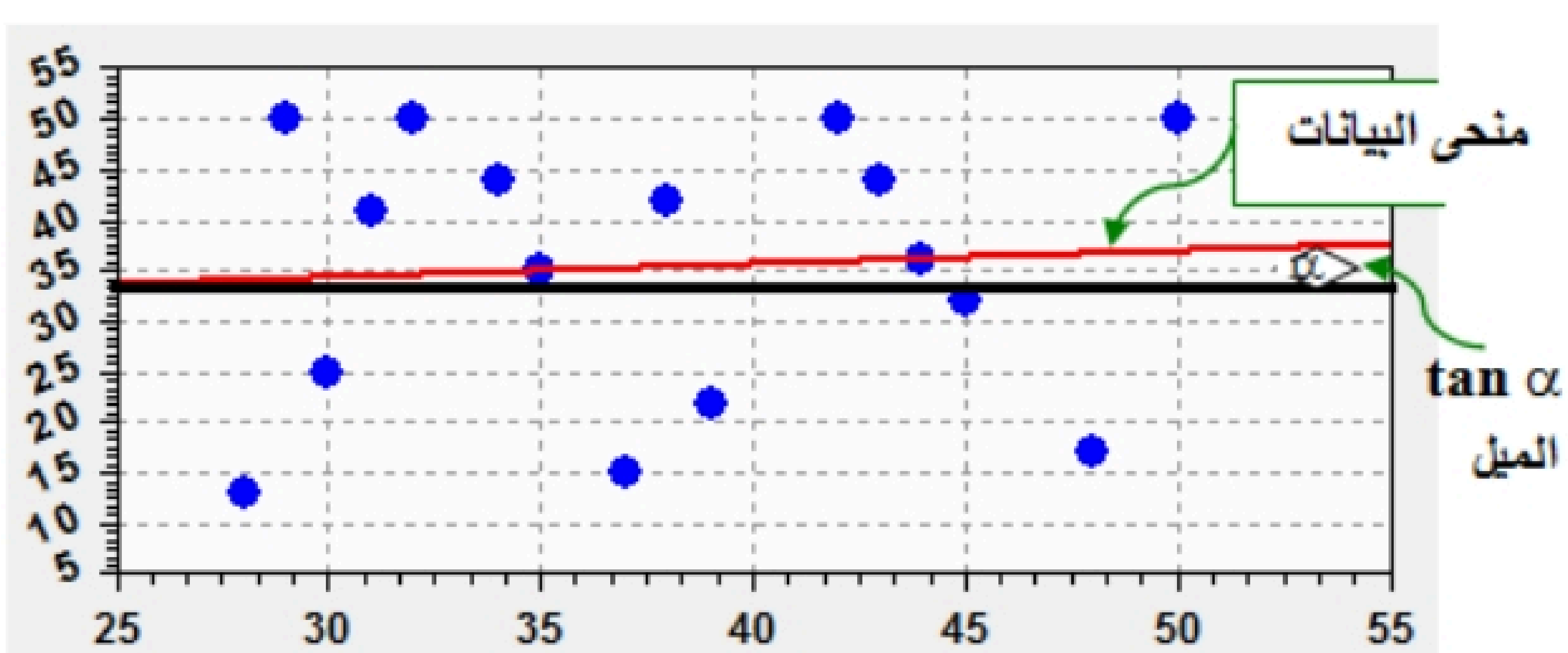
$$S_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})\cdot(y_i-\bar{y}) = \frac{-1868.5}{9} = -207.611$$

وهكذا نجد $S_{X,Y} < 0$ ، وهذا يعني أنَّ مَنْحى البيانات يجب أن يكون هابطاً (وهو موافق لما لاحظناه سابقاً على لوحة الانتشار في الشكل (٣,٢)).

ج- من أجل بيانات المثال / ٣/ من (١-٢-١-٣) لدينا $\bar{x} = 37.813$ وكذلك $\bar{y} = 35.375$ ، ومن ثمَّ سيكون التغاير بين بيانات $X$ و$Y$ يساوي:

$$S_{X,Y} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})\cdot(y_i-\bar{y}) = \frac{95.125}{15} = 6.342$$

وهكذا نجد $S_{X,Y} > 0$ ، وهذا يعني أنَّ مَنْحى البيانات يجب أن يكون صاعداً رغم صعوبة تحديد المنحنى بالعين المجرّدة من خلال لوحة الانتشار، ولكن مَيْل Slope المَنْحى لهذه البيانات سيكون صغيراً (انظر الشكل الآتي، وأمّا تعيين معادلة هذا المَنْحى فسوف نأتي عليها في الفقرة القادمة من هذه الفصل).




الشكل (٦ , ٣)

**(٢ , ٣ , ١ , ٣) مثال**

لتكن لدينا البيانات الآتية الناتجة عن إخضاع مجموعة مكوَّنة من 87 طالباً لاختبارين، أحدهما $X$ يمثّل مقرَّر الرياضيات العامّة (الدرجة القصوى **80**)، والآخر $Y$ يمثّل مقرَّر الإحصاء (الدرجة القصوى **30**).

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| (41,12) | (58,20) | (58,21) | (46,16) | (54,19) | (42,14) | (64,23) | (60,21) | (29,10) | (41,13) | (74,25) |
| (63,22) | (71,25) | (79,29) | (69,24) | (40,12) | (55,19) | (69,24) | (52,19) | (55,20) | (71,25) | (45,15) |
| (49,17) | (74,25) | (31,11) | (53,19) | (38,12) | (49,17) | (36,11) | (49,17) | (40,12) | (46,16) | (61,22) |
| (48,17) | (43,14) | (53,19) | (41,13) | (50,18) | (75,27) | (49,17) | (51,18) | (78,27) | (50,18) | (34,11) |
| (44,15) | (20,10) | (75,26) | (57,20) | (59,21) | (43,14) | (46,16) | (75,27) | (32,11) | (64,23) | (53,19) |
| (50,18) | (69,23) | (23,10) | (67,23) | (47,17) | (79,29) | (79,28) | (61,21) | (52,19) | (64,23) | (39,12) |
| (62,22) | (71,25) | (33,11) | (51,18) | (31,10) | (44,15) | (46,16) | (43,15) | (74,26) | (47,16) | (46,16) |
| (44,15) | (69,24) | (59,21) | (54,19) | (51,18) | (42,14) | (48,17) | (33,11) | (63,23) | (60,21) | |

فلو قمنا بصبِّ هذه البيانات في جدول توزيع تكراري مشترك تكون فيه سعة الفئات للبيانات $X$ تساوي 10، و سعة الفئات للبيانات $Y$ تساوي 4 (على سبيل المثال). فإنّنا سنحصل على جدول تكراري مشترك لهذه البيانات له العرض الآتي:

الشكل (٣,٧)

| فئات $Y$ | فئات $X$ | | | | | | Total |
|---|---|---|---|---|---|---|---|
| | 20→30 | 30→40 | 40→50 | 50→60 | 60→70 | 70→80 | |
| 10 →14 | 3 | 6 | 7 | - | 1 | - | 17 |
| 14 →18 | - | 2 | 11 | 9 | - | - | 22 |
| 18 →22 | - | 1 | 7 | 10 | 5 | - | 23 |
| 22 →26 | - | - | 1 | 1 | 9 | 6 | 17 |
| 26 →30 | - | - | 1 | - | - | 7 | 8 |
| Total | 3 | 9 | 27 | 20 | 15 | 13 | 87 |

وبحساب المقادير المطلوبة لحساب التغاير بين بيانات $X$ و $Y$ نجد أنَّ:

$$S_{X,Y} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}\left(x_i - \overline{x}\right)\left(y_j - \overline{y}\right)}{\left(\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}\right) - 1} = \frac{4462.51}{86} = 51.89$$

الذي يُظهر لنا أنَّ مَنْحى البيانات إيجابي ، وأمّا بشأن قوة التّأثير المتبادل بين المقرَّرين فإنَّه بحاجة لدراسة أخرى سنأتي على ذكرها لاحقاً.

لقد اقتُرِحَ أَخْذ التغاير بين البيانات كمقياس لمدى الارتباط بين هذه البيانات، ولكن لوحظ أنّه من الممكن أن تكون وحدة القياس لكل من القيم $x_i$ و $y_i$ مختلف بعضها عن بعض (مثل: الطول بـ المتر، والوزن بـ الكيلوغرام)، ومن ثَمَّ يصبح استخدام التغاير لقياس التّأثير المتبادل بين البيانات عديم الجدوى، ولهذا قدّمت بعض المقاييس التي تُعطي دلالات أفضل حول التّأثير المتبادل بين البيانات ولا تحتوي على التناقضات التي وجدت في التغاير، ومن هذه المقاييس نقدّم المقاييس الآتية.

### (٣ , ١ , ٤) مُعامل بيرسون للارتباط Person's Coefficient of Correlation:

لقد لاحظ بيرسون أنَّ قيمة التغاير بين البيانات مرتبطة بوحـدة القياس، وعلاوةً على ذلك لاحظ أنَّ اختلاف قيمة التشتت لكلٍّ من مجموعتي البيانات قد يؤثِّر في قيمة هذا المقياس أيضاً، ولذلك قام بتقسيم قيمة التغاير على حاصل ضرب الانحراف المعياري لكل من مجموعتي البيانات، فحصل نتيجة لذلك على مقدار عددي مجرَّد من وحدة القياس للبيانات ولا يتأثَّر بتشتت البيانات. لذلك قدَّم هذا المقياس لاستخدامه في قياس قوة الارتباط الخطي بين مجموعتين من البيانات . أمّا لحساب مُعامل بيرسون للارتباط الخطي (وسنرمز له بـ $\rho_P$ من أجل المجتمع وبـ $\mathbf{r}_P$ من أجل العيّنة) فإنَّنا سنناقش الحالتين الآتيتين:

**أولاً: إذا كانت البيانات المُعطاة خام:**

١ - بفرض أنَّ $(x_1,y_1)$ ، $(x_2,y_2)$ ،... و $(x_N,y_N)$ بيانات مجتمع خاضع لمتغيِّرين $X$ و $Y$، وأنَّ للبيانات $x_1$ ، $x_2$ ،... و $x_N$ متوسط وانحراف معياري $\mu_X$ و $\sigma_X$ على الترتيب، وكذلك للبيانات $y_1$ ، $y_2$ ،... و $y_N$ متوسط وانحراف معياري $\mu_Y$ و $\sigma_Y$ على الترتيب، فعندئذ يُحسب مُعامل بيرسون للارتباط الخطي لهذه البيانات $\rho_P$ بالعلاقة الآتية:

$$\rho_P := \begin{cases} \dfrac{\sigma_{X,Y}}{\sigma_X \cdot \sigma_Y} & for\ \sigma_X , \sigma_Y > 0 \\ 0 & eleswhere \end{cases} \qquad [3,4\text{-}a]$$

ومن أجل $\sigma_X$ و $0 < \sigma_Y$ يمكن استخدام إحدى العلاقات المكافئة الآتية لحسابه أيضاً:

$$\rho_P = \frac{1}{N}\sum_{i=1}^{N} \frac{(x_i - \mu_X)}{\sigma_X} \cdot \frac{(y_i - \mu_Y)}{\sigma_Y} \qquad [3,4\text{-b}]$$

$$\rho_P = \frac{\sum_{i=1}^{N} x_i y_i - N\ \mu_X\ \mu_Y}{\sqrt{\sum_{i=1}^{N} x_i^2 - N\ \mu_X^2} \cdot \sqrt{\sum_{i=1}^{N} y_i^2 - N\ \mu_Y^2}} \qquad [3,4\text{-c}]$$

$$\rho_P = \frac{N\left(\sum_{i=1}^{N} x_i y_i\right) - \left(\sum_{i=1}^{N} x_i\right)\cdot\left(\sum_{i=1}^{N} y_i\right)}{\sqrt{N\left(\sum_{i=1}^{N} x_i^2\right) - \left(\sum_{i=1}^{N} x_i\right)^2} \cdot \sqrt{N\left(\sum_{i=1}^{N} y_i^2\right) - \left(\sum_{i=1}^{N} y_i\right)^2}} \qquad [3,4\text{-d}]$$

٢- بفرض أنَّ $(x_1, y_1)$ ، $(x_2, y_2)$ ،...و $(x_n, y_n)$ بيانات عيِّنة أخضعت لمتغيِّرين $X$ و $Y$ ، وأنَّ للبيانات $x_1$ ، $x_2$ ،... و $x_n$ متوسط وانحراف معياري $\overline{x}$ و $S_X$ على الترتيب، وكذلك للبيانات $y_1$ ، $y_2$ ،... و $y_n$ متوسط وانحراف معياري $\overline{y}$ و $S_Y$ على الترتيب، فعندئذ يُحسب مُعامل بيرسون للارتباط الخطي بين هذه البيانات بالعلاقة الآتية:

$$\mathbf{r}_P := \begin{cases} \dfrac{S_{X,Y}}{S_X \cdot S_Y} & \textit{for } S_X, S_Y > 0 \\ 0 & \textit{eleswhere} \end{cases} \qquad [3,5\text{-a}]$$

ومن أجل $S_X$ و $0 < S_Y$ يمكن استخدام إحدى العلاقات المكافئة الآتية لحسابه:

$$\mathbf{r}_P = \frac{1}{n-1}\sum_{i=1}^{n} \frac{(x_i - \overline{x})}{S_X} \cdot \frac{(y_i - \overline{y})}{S_Y} \qquad [3,5\text{-b}]$$

$$\mathbf{r}_P = \frac{\sum_{i=1}^{n} x_i\ y_i - n\ \overline{x}\ \overline{y}}{\sqrt{\sum_{i=1}^{n} x_i^2 - n\ \overline{x}^2} \cdot \sqrt{\sum_{i=1}^{n} y_i^2 - n\ \overline{y}^2}} \qquad [3,5\text{-c}]$$

$$\mathbf{r}_P = \frac{n\cdot\left(\sum_{i=1}^{n} x_i y_i\right) - \left(\sum_{i=1}^{n} x_i\right)\cdot\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{n\cdot\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n\cdot\left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} y_i\right)^2}} \qquad [3,5\text{-d}]$$

**ثانياً: إذا كانت البيانات المُعطاة مُجمَّعة في جدول تكراري مشترك:**

بفرض أنَّه لدينا بيانات مُعطاة من خلال جدول تكراري مشترك كما في الجدول (٦ ,٣)، فعندئذ:

١- إذا كانت هذه البيانات لمجتمع حجمه منته ويساوي $N$ وأخضع لمتغيِّرين $X$ و $Y$ ، ومن ثمَّ صُبَّت نتائجها في ذلك الجدول، وكان الانحراف المعياري لبيانات $X$ و $Y$ هو $\sigma_X$ و $\sigma_Y$ على الترتيب مع $\sigma_X, \sigma_Y > 0$ ، فعندئذ يُحسب مُعامل بيرسون للارتباط الخطي لبيانات ذلك الجدول من خلال العلاقة الآتية:

$$\rho_P = \frac{\sigma_{X,Y}}{\sigma_X \cdot \sigma_Y} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{\ell} f_{ij} \left(x_i - \mu_X\right) \left(y_j - \mu_Y\right)}{\sigma_X \cdot \sigma_Y \cdot \sum_{i=1}^{k} \sum_{j=1}^{\ell} f_{ij}} \qquad [3,6\text{-a}]$$

علماً أنَّ $x_i$ هو مركز الفئة $\alpha_{i+1} \to \alpha_i$ من أجل كل $i \ni \mathbb{N}_k$، وبالمثل $y_j$ هو مركز الفئة $\beta_{j+1} \to \beta_j$ من أجل كل $j \ni \mathbb{N}_\ell$، وأمّا $f_{ij}$ فهو التكرار المشترك بين الفئتين $(\alpha_i , \alpha_{i+1}]$ و $(\beta_j , \beta_{j+1}]$.

أمّا إذا كان $\sigma_X = 0$ أو $\sigma_Y = 0$ ("أو" هنا لا تفيد الحصر) فإنَّنا سنضع $\rho_P = 0$. علماً أنَّ $\sigma_X^2$ و $\sigma_Y^2$ يحسبان كما يلي:

$$\sigma_X^2 = \frac{1}{\sum_{i=1}^{k} \sum_{j=1}^{\ell} f_{ij}} \sum_{i=1}^{k} \left( \left(x_i - \overline{x}\right)^2 \cdot \sum_{j=1}^{\ell} f_{ij} \right) \qquad [3,6\text{-b}]$$

$$\sigma_Y^2 = \frac{1}{\sum_{i=1}^{k} \sum_{j=1}^{\ell} f_{ij}} \sum_{j=1}^{\ell} \left( \left(y_j - \overline{y}\right)^2 \cdot \sum_{i=1}^{k} f_{ij} \right) \qquad [3,6\text{-c}]$$

٢- إذا كانت هذه البيانات لعيِّنة حجمها $n$ أخضعت لمتغيِّرين $X$ و$Y$، ومن ثمَّ صُبَّت نتائجها في الجدول (٦ , ٣)، وبفرض أنَّ الانحراف المعياري لبيانات $X$ و$Y$ هو $S_X$ و $S_Y$ على الترتيب مع $S_X$ و $S_Y > 0$ ، فعندئذ يُحسب مُعامل بيرسون للارتباط الخطي لبيانات ذلك الجدول من خلال العلاقة الآتية:

$$\mathbf{r}_P = \frac{S_{X,Y}}{S_X \, S_Y} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{\ell} f_{ij} \left(x_i - \overline{x}\right) \left(y_j - \overline{y}\right)}{S_X \, S_Y \left( \sum_{i=1}^{k} \sum_{j=1}^{\ell} f_{ij} - 1 \right)} \qquad [3,7\text{-a}]$$

وأمّا إذا كان $S_X = 0$ أو $S_Y = 0$ فإنَّنا نضع $\mathbf{r}_P = 0$ . علماً أنَّ $S_X^2$ و $S_Y^2$ يُحسبان كما يلي:

$$S_X^2 = \frac{1}{\left( \sum_{i=1}^{k} \sum_{j=1}^{\ell} f_{ij} \right) - 1} \sum_{i=1}^{k} \left( \left(x_i - \overline{x}\right)^2 \cdot \sum_{j=1}^{\ell} f_{ij} \right) \qquad [3,7\text{-b}]$$

$$S_Y^2 = \frac{1}{\left( \sum_{i=1}^{k} \sum_{j=1}^{\ell} f_{ij} \right) - 1} \sum_{j=1}^{\ell} \left( \left(y_j - \overline{y}\right)^2 \cdot \sum_{i=1}^{k} f_{ij} \right) \qquad [3,7\text{-c}]$$

وهنا نودّ أن نشير إلى الأمور الآتية:

أ- يُستخدم هذا المقياس من أجل البيانات الكميّة فقط، ويُعدّ من المقاييس الجيدة للارتباط الخطي، وجودته تكمن في أنَّه يستخدم جميع قيم بيانات المتغيِّرين $X$ و$Y$ في الحساب.

ب- إنَّ الصيغ المستخدمة لحساب مُعامل الارتباط الخطي لـ بيرسون هي نفسها لجميع العلاقات السّابقة، ولكن عُدِّل عليها لكي تتوافق مع المُعطيات المتوفرة من البيانات.

ج) عندما يُطلب من برنامج إحصائي (مثل Minitab أو SPSS) حساب مُعامل بيرسون للارتباط الخطي للبيانات المقدّمة فإنَّه يتعامل معها على أنَّها بيانات عيِّنة فقط.

من المعلوم أنَّ البيانات التي نحصل عليها نتيجة دراسة إحصائية على مجتمع أو عيِّنة قد تكون نوعيّة أو كميّة أو مختلطة (بمعنى أنَّ إحداهما كميّة والأخرى نوعيّة) في حال البيانات المتزاوجة، وهذا يعني أنَّه إذا كانت البيانات الواجب دراسة الارتباط بينها هي بيانات نوعيّة أو مختلطة فإنَّ مُعامل بيرسون الذي قُدِّم قبل قليل لم يَعُد صالحاً للاستخدام، ومن ثَّم يجب تقديم مقياس آخر لإيجاد مُعامل الارتباط بحيث يعطي دلالة جيدة لمدى التَّأثير المتبادل بين مجموعتي البيانات. إنَّ الفقرة الآتية تقدّم لنا مقياساً مقبولاً لتعيين نوع ودرجة الارتباط لبيانات نوعيّة، أو كميّة، أو مختلطة.

### (٣ , ١ , ٥) مُعامل سبيرمان لارتباط الرتب Spearman's Rank Correlation Coefficient

لنفترض أنَّه لدينا عيِّنة مسحوبة من مجتمع إحصائي، ونتائجها ثنائيات ( أي كل عنصر أُخضع لدراسة متغيِّرين أو ظاهرتين $X$ و$Y$)، وأنَّ بيانات هذه العيِّنة هي بيانات كميّة، نوعيّة أو مختلطة، ولنقم ببناء جدول يُدوَّن في صفّه العلوي عناصر العيِّنة أو رموزها (انظر الجدول الآتي وتابع الخطوات القادمة عليه)، ومن ثَّم نقوم بتنفيذ الخطوات الآتية:

الشكل (٨ , ٣)

| العنصر | $e_1$ | $e_2$ | ..... | $e_{n-1}$ | $e_n$ |
|---|---|---|---|---|---|
| قيم المتغيِّر $X$ | $x_1$ | $x_2$ | ..... | $x_{n-1}$ | $x_n$ |
| الرتب الأولية $\alpha_i$ | $\alpha_1$ | $\alpha_2$ | ..... | $\alpha_{n-1}$ | $\alpha_n$ |
| الرتب النهائية $\alpha_i^*$ | $\alpha_1^*$ | $\alpha_2^*$ | ..... | $\alpha_{n-1}^*$ | $\alpha_n^*$ |
| قيم المتغيِّر $Y$ | $y_1$ | $y_2$ | ..... | $y_{n-1}$ | $y_n$ |
| الرتب الأولية $\beta_i$ | $\beta_1$ | $\beta_2$ | ..... | $\beta_{n-1}$ | $\beta_n$ |
| الرتب النهائية $\beta_i^*$ | $\beta_1^*$ | $\beta_2^*$ | ..... | $\beta_{n-1}^*$ | $\beta_n^*$ |
| $d_i = \alpha_i^* - \beta_i^*$ | $d_1$ | $d_2$ | ..... | $d_{n-1}$ | $d_n$ |
| $d_i^2$ | $d_1^2$ | $d_2^2$ | ..... | $d_{n-1}^2$ | $d_n^2$ |

١- إعطاء كل بيان $x_i$ و $y_i$ من هذه البيانات رقماً ندعوه الرتبة الأوليّة Initial Rank وسنرمز له بـ $\alpha_i$ و $\beta_i$ لكل $i \in \mathbb{N}_n$ على الترتيب، وبحيث يكون لهذه الرتب الأوليّة صفة تصاعديّة (أو تنازليّة) ولكلٍّ من مجموعتي البيانات $X$ و$Y$ بآن واحد، بمعنى أنَّه لا يجوز أنَّ نعطي إحدى مجموعتي البيانات رتباً تصاعديّة والأخرى رتباً تنازليّة.

٢- تعيين المتوسط لرتب البيانات المتماثلة (أو القيم المتساوية في حالة البيانات الكميّة)، ومن ثَّم إعطاء كل واحدة من هذه البيانات المتماثلة (أو القيم المتساوية) قيمة هذا المتوسط التي تُدعى الرتبة النهائية Final Rank، وسنرمز لها بـ $\alpha_i^*$ و $\beta_i^*$ لكل $i \in \mathbb{N}_n$ .

٣- حساب الفروقات بين الرتب النهائية المتقابلة $\alpha_i^*$ و $\beta_i^*$، وسنرمز لهذه الفروق بـ $d_i$ لكل $i \in \mathbb{N}_n$، أي إن:

$$d_i = \alpha_i^* - \beta_i^* \qquad [3,8]$$

٤- نقوم بتربيع قيم الفروقات $d_i$ فنحصل على $d_i^2$ لكل $i \in \mathbb{N}_n$ .

٥- نضع ما يلي:

$$\mathbf{r}_S := 1 - \frac{6 \sum_{i=1}^{n} d_i^2}{n\,(n^2-1)} \qquad [3,9]$$

إنَّ قيمة $\mathbf{r}_S$ التي اشتقها سبيرمان كحالة خاصة من علاقة بيرسون العزومية للارتباط لـدى استخدام رتب القيم أو القراءات بدلاً من القيم نفسها تُدعى **مُعامل سبيرمان لارتباط الرتب** (**أو مُعامل الارتباط الرتبي لـ سبيرمان**)، وهذه القيمة تقيس قوة الارتباط الخطي بين ظاهرتين. إنَّ العلاقة [3,9] ترجع إلى الإحصائي البريطاني **سبيرمان** **Charles Edward Spearman (1863-1945)**، كما يطلق عليها اسم **مُعامل سبيرمان- رو للارتباط** Spearman- Rho Correlation Coefficient أيضاً.

### (١ , ٥ , ١ , ٣) ملاحظات

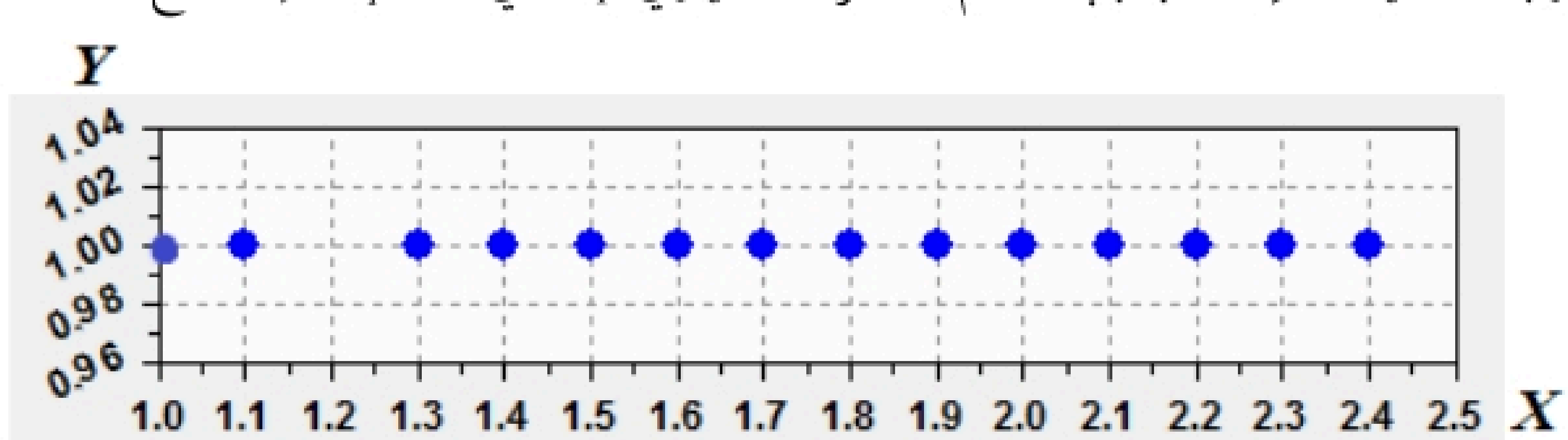
١- إنَّ مُعامل الارتباط الخطي لبيانات عيِّنة يُعدّ مُقدِّراً لمعامل الارتباط الخطي لبيانات المجتمع، ولكن عند استخدام البرامج الإحصائية لحساب مُعامل الارتباط الخطي لبيانات متغيِّرين $X$ و$Y$، فإنَّ معظم هذه البرامج تتعامل مع البيانات على أنَّها بيانات عيِّنة فقط، ولذلك عندما نكتب $\mathbf{r}$ فإنَّنا سـ نقصد بذلك أي مُعامل من مُعاملات الارتباط الخطي سواء كانت البيانات لعيّنات أو لمجتمعات.

٢- إنَّ قيمة $\mathbf{r}$ تنتمي إلى الفترة $[-1 , +1]$ دوماً، وإذا كان:

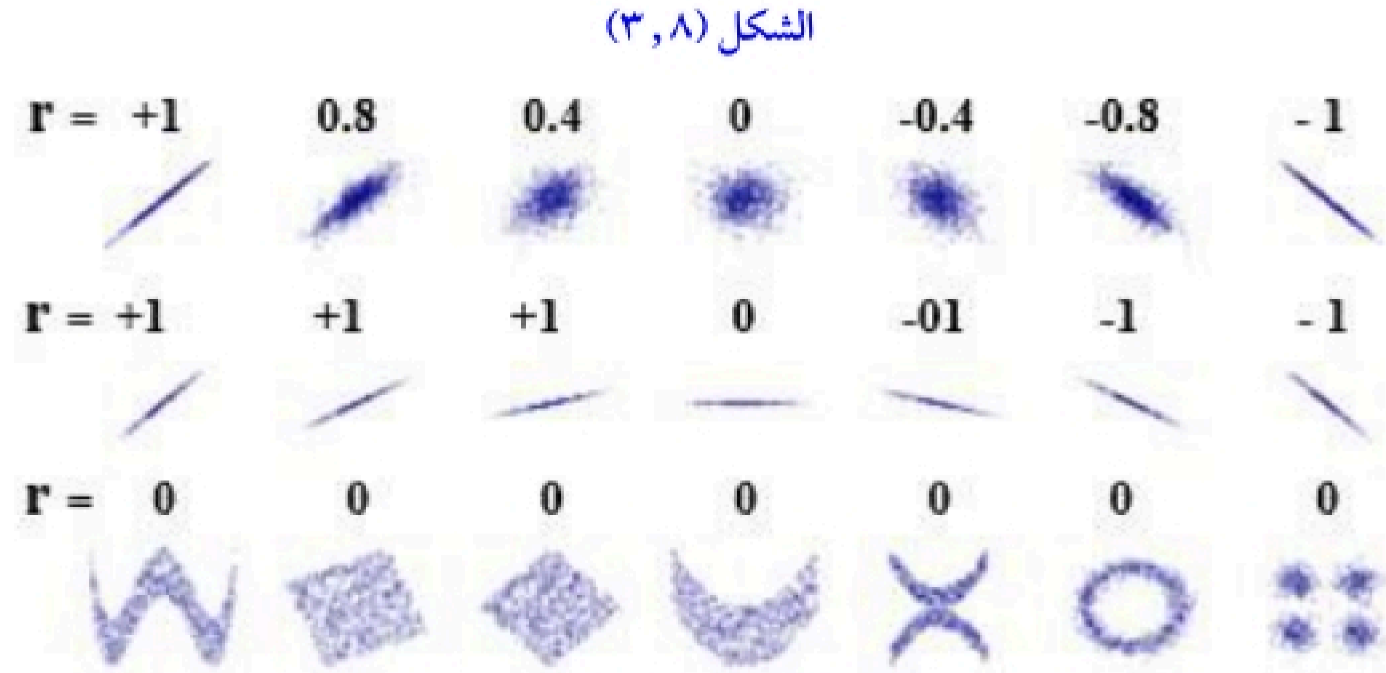
أ - لدينا $\mathbf{r} > 0$ فعندئذ يُقال إنَّ الارتباط بين المتغيِّرين $X$ و$Y$ **إيجابي** (**أو طردي**).

ب- لدينا $\mathbf{r} < 0$ فإنَّه يُقال إنَّ الارتباط بين المتغيِّرين $X$ و$Y$ **سلبي** (**أو عكسي**).

ج- لدينا $\mathbf{r} = \pm 1$ فعندئذ يُقال عن الارتباط بين المتغيِّرين $X$ و$Y$ إنَّه خطي تام، وفي هذه الحالة تكون جميع النقاط الممثِّلة للمشاهدات $(x_i , y_i)$ مع $i \in \mathbb{N}_n$ تقع على خط مستقيم واحد. لكن يجب التنبيه هنا إلى أنَّ العكس ليس صحيحاً بالضرورة، فلو نظرنا إلى مجموعة البيانات الموضَّحة من خلال الشكل الآتي (٣ , ٨) لوجدنا أنَّ جميع النقاط الممثِّلة لها تقع على مستقيم واحد، ومن ثمَّ ترتبط مع بعضها ارتباطاً خطياً تاماً، إلاَّ أنَّه بسبب انعدام الانحراف المعياري $S_Y$ (أي $S_Y = 0$) فإنَّنا نضع $\mathbf{r} = 0$ .




الشكل (٣ , ٨)




الشكل (٣ , ٩)

د - لدينا $\mathbf{r} = 0$ فعندئذ **لا** يُقال إنَّه **لا يوجد ارتباط** بين الظاهرتين، بل نكون أمام إحدى الحالات الثلاث الآتية:

د-١- توزّع البيانات على لوحة انتشار له شكل غيمة عديمة الاتجاه، أو في تجمعات جزئية متناثرة تشكِّل بمجملها مشهداً يصعب تحديد اتجاهه، ومن ثمَّ مَنْحى البيانات سيكون غير واضح، ومن الأمثلة على ذلك الشكل (٣ , ٤) حيث لدينا $S_{X,Y} = 0$، ومن ثمَّ $\mathbf{r} = 0$، وكذلك بعض العروض التوضيحية الأخرى تجدها في الشكل الآتي (٣ , ٩).

د-٢- مَنْحى البيانات يوازي أحد محوري لوحة الانتشار، ومن الأمثلة على ذلك الشكل (٥ , ٣) حيث لدينا $S_{X,Y} = 0$ ومن ثمَّ $\mathbf{r} = 0$، وكذلك بعض العروض التوضيحيّة الأخرى تجدها في الشكل الآتي (٩ , ٣).

د-٣- توزّع البيانات على لوحة انتشار له شكل بيان علاقة غير خطيّة بين المتغيِّرين $X$ و$Y$، ونماذج توضيحية على ذلك تجدها في الشكل الآتي (٩ , ٣).

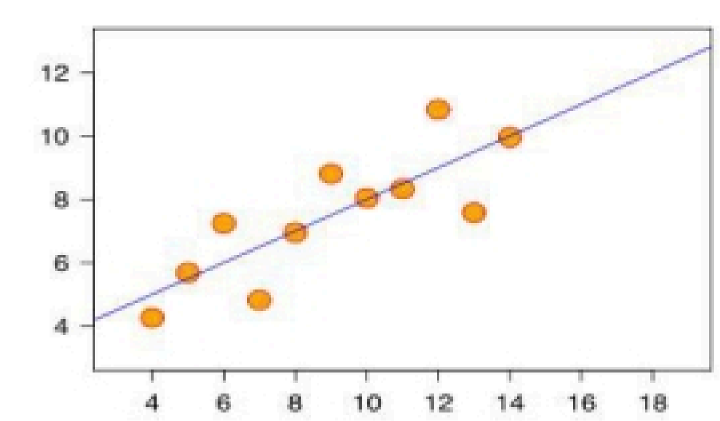
٣- لتقييم **قوة** و**ضعف** الارتباط الخطي بين بيانات المتغيِّرين $X$ و$Y$، فإنّه يُقال إنَّ هذا الارتباط:

الشكل (٩ , ٣)

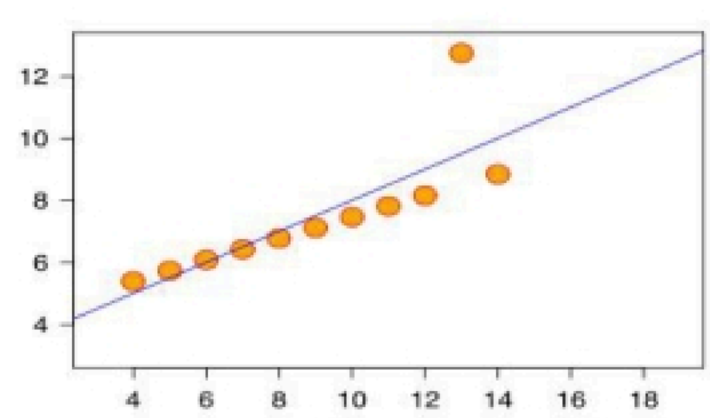
| قيمة معامل الارتباط | $0.86 < \lvert\mathbf{r}\rvert \leq 1$ | $0.70 < \lvert\mathbf{r}\rvert \leq 0.86$ | $0.50 < \lvert\mathbf{r}\rvert \leq 0.70$ | $0.30 < \lvert\mathbf{r}\rvert \leq 0.50$ | $0 \leq \lvert\mathbf{r}\rvert \leq 0.30$ |
|---|---|---|---|---|---|
| الحكم على الارتباط | قوي جداً | قوي | متوسط القوة | ضعيف | ضعيف جداً |

وهذه المناقشة تبقى سارية المفعول على كافة مقاييس الارتباط الخطي للبيانات سواء كانت لعيِّنات أو لمجتمعات.

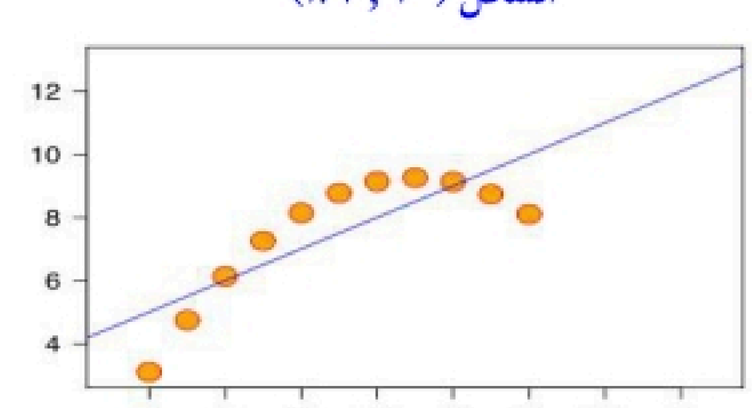
٤- من الممكن أن يكون لدينا مجموعات بيانات (على شكل ثنائيات) ذات توزيعات مختلف بعضها عن بعض الآخر، ومع ذلك تمتلك قيمة واحدة لمعامل الارتباط الخطي، فلو أمعنا النظر في سلسلة العروض البيانية الآتية التي تُشتهر باسم **رباعية إنسكومب** Anscombe's Quartet، وكان قد قدَّمها الإحصائي الإنكليزي **إنسكومب** Francis John Anscombe (1918-2001) عام (1973) في جامعة كامبريدج.



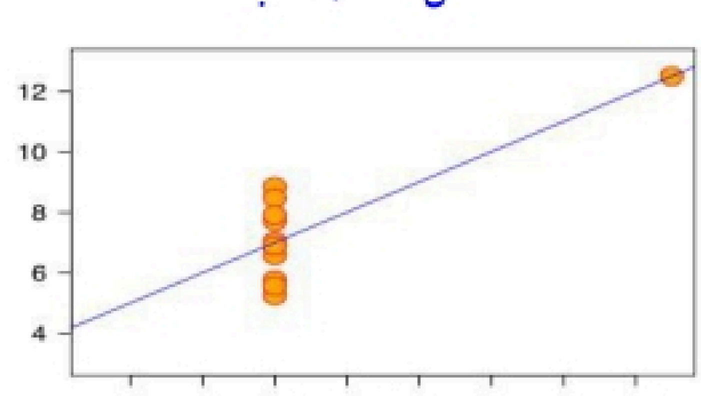
الشكل (١٠ , ٣ .أ)



الشكل (١٠ , ٣ .ب)



الشكل (١٠ , ٣ .ج)



شكل (١٠ , ٣ .د)

وهذه العروض تذكرنا بأنّه في كثير من الأحيان يكون من المهمّ جداً النظر إلى توزّع النقاط الممثِّلة للبيانات على لوحة الانتشار قبل تحليلها، حيث يمكن للقيم المتطرِّفة وأنماط أخرى من التوزّع غير الخطي للبيانات أن يكون لها تأثير كبير في دراسة البيانات وعلى النتائج المتمخِّضة عنها، ففي الأشكال السّابقة لدينا أربع مجموعات من البيانات المتزاوجة لها الخصائص الآتية نفسها:

المتوسط $(\overline{x}, \overline{y}) = (7, 5)$، التباين $(S_X^2, S_Y^2) = (4, 2)$، معامل ارتباط $\mathbf{r} = 0.816$ وخط الانحدار $\hat{Y} = 3 + 0.5X$ (سنأتي على ذكر هذا المفهوم لاحقاً).

لكن كما هو ملاحظ من لوحات الانتشار الأربع أنَّ توزيع المتغيِّرات مختلفة جداً من مجوعة إلى أخرى. لا بل اثنتين منهما (اللوحة في الأسفل (٨ , ٣ج) و (٨ , ٣د) تخضع لعلاقة غير خطيّة، وهذه الأمثلة تظهر لنا بجلاء أنَّ مُعامل الارتباط لا يشير بالضبط إلى أن هناك علاقة داليّة خطيّة بين $X$ و $Y$، وإنّما يشير إلى أي مدى يمكن لهذه العلاقة أن تقترب من علاقة خطيّة فقط.

٥- لقد لاحظنا أنَّ قيمة مُعامل الارتباط الخطي **r** تشير إلى قوة وجود علاقة خطيّة بين متغيِّرين $X$ و$Y$، ولكن هذه القيمة لا تميّز العلاقة بينهما تماماً في الحالة العامّة، وعلى وجه الخصوص إذا كان التوقع الشَرطي لـ $Y$ بالنسبة إلى $X$ ليست خطيّة في $X$ (للاطلاع على مفهوم التوقع الشرطي للمتغيِّرات العشوائية يمكن الرجوع إلى نهاية الفصل السابع).

٦- ربما يكون الارتباط من أكثر المفاهيم تعاطيا في الإحصاء. حيث نلاحظ أنَّ كثيراً من الناس يستخدمون كلمة **ارتباط** على أنَّها تعني أي نوع من الارتباط بين متغيِّرين $X$ و$Y$، ولكن في الحقيقية له معنى تقني صارم للغاية، إذ إنَّه يعني القوةلعلاقة خطيّة واضحة بين متغيِّرين $X$ و$Y$ على مجال $\mathbb{R}^2 \supseteq I_1 \times I_2$.

### أمثلة (٣, ٥, ١, ٢)

١- بالرجوع إلى الأمثلة (٣, ١, ٢, ١) نجد ما يلي:

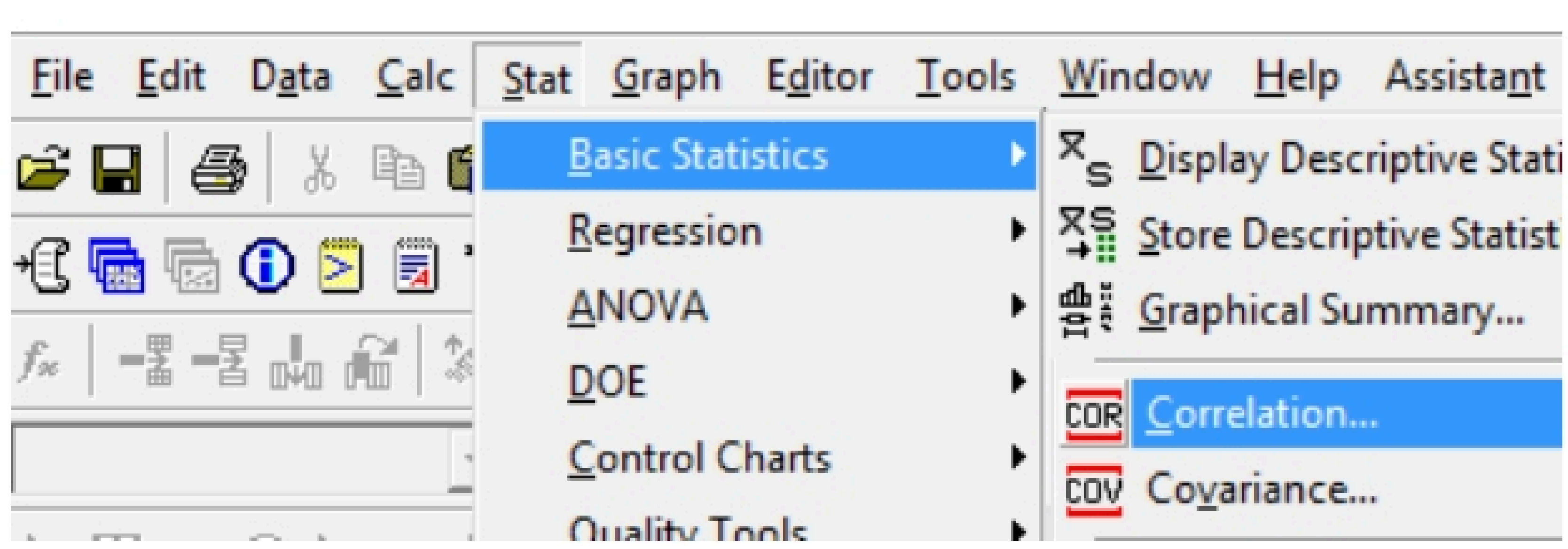
أ- من أجل بيانات المثال / ١/ من (٣, ١, ٢, ١) نجد أنَّ:

$$\overline{x} = 75.9 \quad \& \quad S_X = 12.467 \quad \& \quad \overline{y} = 72.6 \quad \& \quad S_Y = 16.029 \quad \& \quad S_{X,Y} = 5740804$$

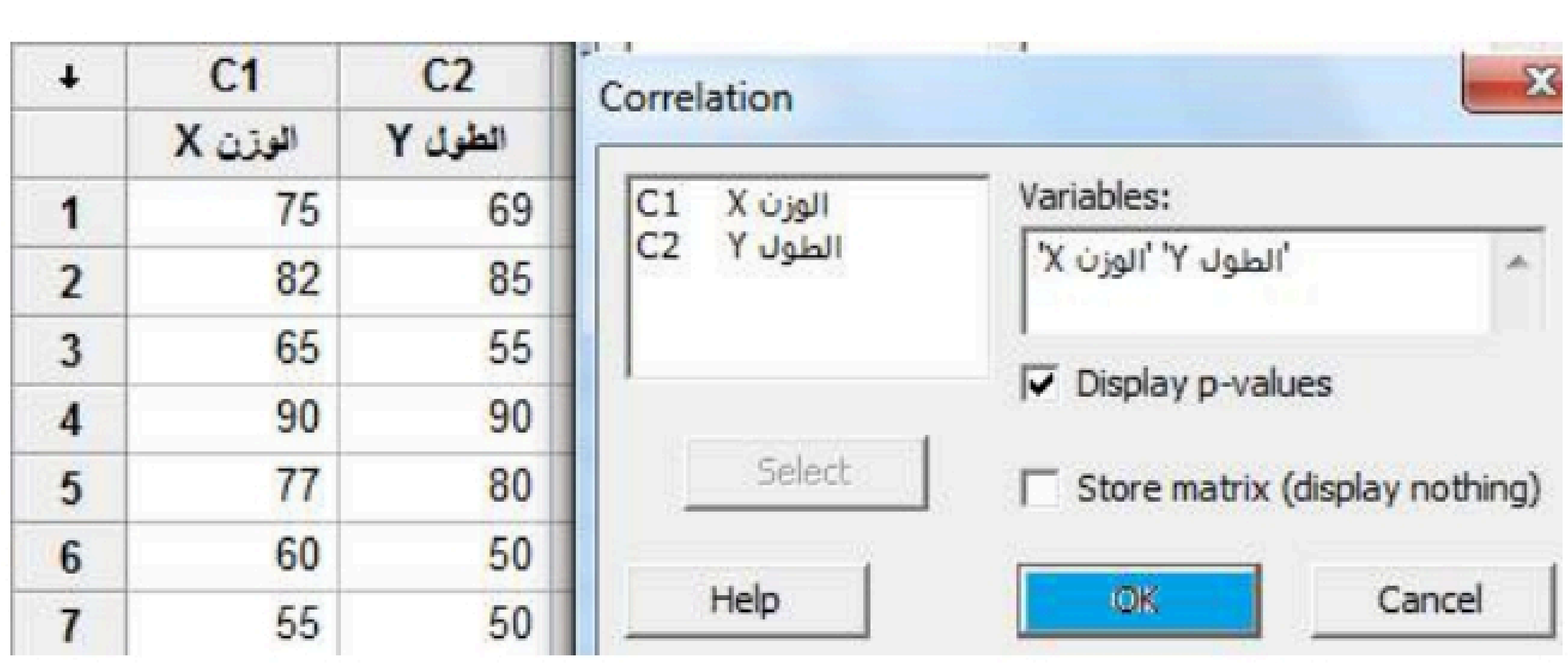
ومن ثمَّ تكون قيمة مُعامل بيرسون للارتباط الخطي بين بيانات $X$ و$Y$ تساوي:

$$\mathbf{r}_P = \frac{S_{X,Y}}{S_X \cdot S_Y} = \frac{5740804}{12.467 \times 16.029} = 0.965$$

فنلاحظ هنا إنَّ الارتباط الخطي بين البيانات $X$ و$Y$ إيجابي قوي جداً لأنَّ قيمته قريبة جداً من $+1$ (انظر إلى مَنْحى هذه البيانات في الشكل (٣, ١ .ب)). كما يمكن استخدام برنامج Minitab لحساب مُعامل بيرسون للارتباط الخطي بين بيانات $X$ و$Y$ وفقاً للخطوات الآتية:




ومن ثمَّ




فنحصل على التقرير الآتي:

Correlations: X الوزن ; Y الطول
Pearson correlation of X and Y = 0.965

ب- من أجل بيانات المثال / ٢/ من (١ , ٢ , ١ , ٣) نجد ما يلي:

$$\overline{x} = 53.1 \quad \& \quad S_X = 21.005$$

$$\overline{y} = 120.5 \quad \& \quad S_Y = 10.080 \quad \& \quad S_{X,Y} = -207.611$$

ومن ثمَّ تكون قيمة $\mathbf{r}_P$ لبيانات $X$ و$Y$ هي:

$$\mathbf{r}_P = \frac{S_{X,Y}}{S_X \cdot S_Y} = \frac{-207.611}{21.467 \times 10.080} = -0.959$$

وكما نلاحظ أنَّ الارتباط الخطي بين البيانات $X$ و$Y$ سلبي وقوي جداً لأنَّ قيمته قريبة جداً من 1− (انظر إلى مَنْحى هذه البيانات في الشكل (٢ , ٣ .ب)).

ج- من أجل بيانات المثال / ٣/ من (١ , ٢ , ١ , ٣) نجد القيم الآتية:

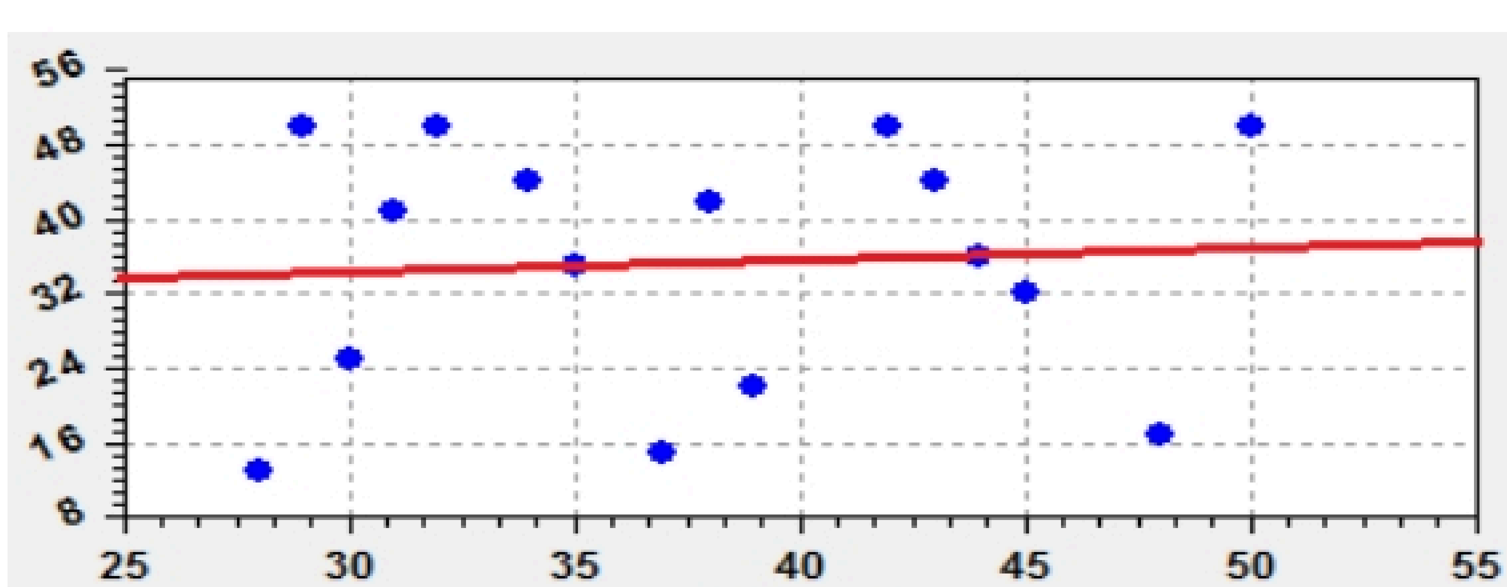
$$\overline{x} = 37.813 \quad \& \quad S_X = 6.9595$$

$$\overline{y} = 35.375 \quad \& \quad S_Y = 13.246 \quad \& \quad S_{X,Y} = 6.342$$

ومنه ينتج لدينا أنَّ قيمة $\mathbf{r}_P$ لبيانات $X$ و$Y$ تساوي:

$$\mathbf{r}_P = \frac{S_{X,Y}}{S_X \cdot S_Y} = \frac{6.342}{6.959 \times 13.246} = 0.069$$

فنلاحظ أنَّ الارتباط الخطي بين بيانات $X$ و$Y$ إيجابي وضعيف جداً لأنَّ قيمته قريبة من الصفر (انظر إلى مَنْحى هذه البيانات في الشكل الآتي).




الشكل (١١ , ٣)

كما أنَّنا نلاحظ أنَّ مَنْحى النقاط الممثِّلة للبيانات يكاد يكون أفقياً، ولكن هذا لا يعني أنَّ صِغَر ميل هذا المَنْحى هو السبب في ضعف الارتباط، وهذا ما أشرنا إليه سابقاً.

٢- بالرجوع إلى المثال (٢ , ٣ , ١ , ٣) والمحمَّلة في الجدول (٥ , ٣)، فإنَّه بعد إجراء العمليات الحسابية لبعض المقادير المطلوبة من أجل حساب مُعامل بيرسون للارتباط الخطي نجد ما يلي:

$$\sum_{i=1}^{k} \sum_{j=1}^{\ell} f_{ij} \left(x_i - \overline{x}\right) \left(y_j - \overline{y}\right) = 4462.51 \quad \& \quad S_X = 13.350 \quad \& \quad S_Y = 4.943 \quad \& \quad \sum_{i=1}^{k} \sum_{j=1}^{\ell} f_{ij} = 87$$

ومن ثمَّ تكون قيمة مُعامل بيرسون للارتباط الخطي هي:

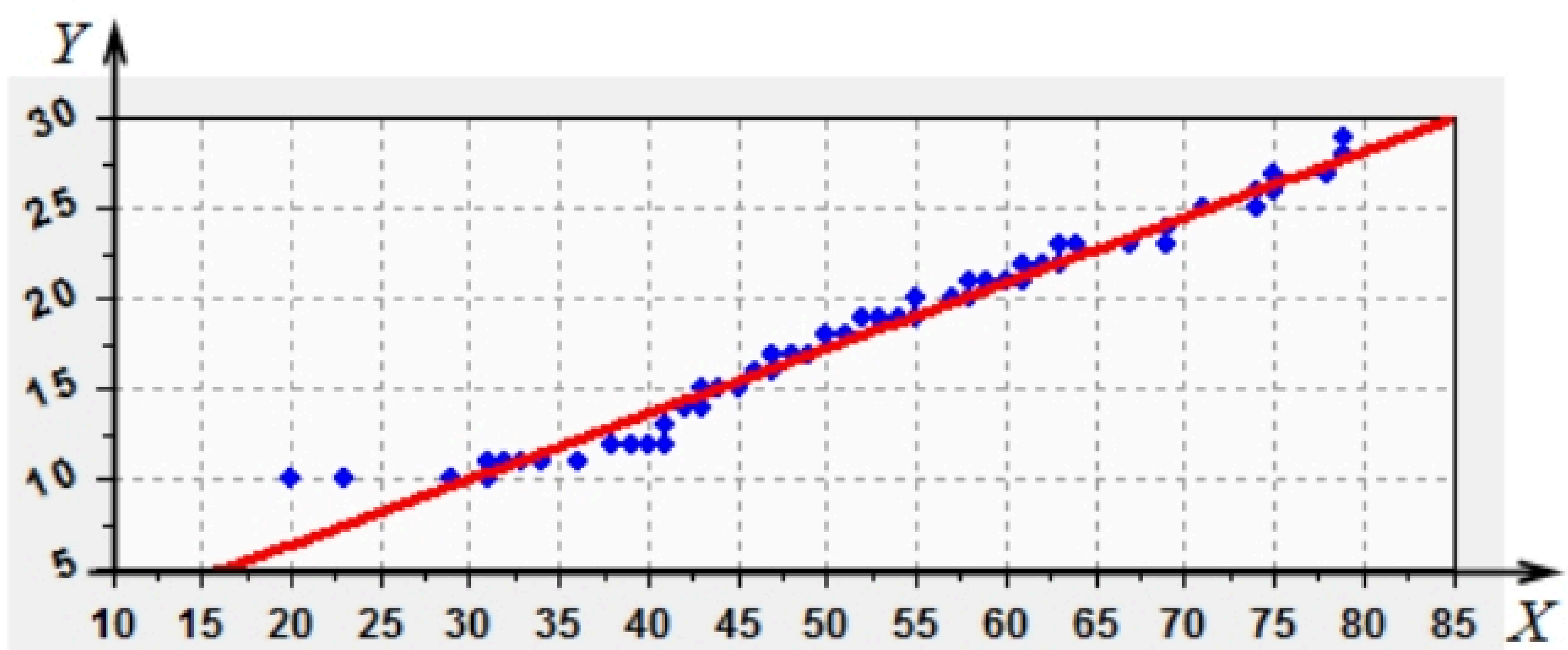
$$\mathbf{r}_P=\frac{\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}\left(x_i-\overline{x}\right)\left(y_j-\overline{y}\right)}{S_X\,S_Y\left(\sum_{i=1}^{k}\sum_{j=1}^{\ell} f_{ij}-1\right)}=\frac{4462.51}{(13.350)\cdot(4.943)\cdot(86)}=0.786$$

حيث نلاحظ أنَّ الارتباط بين بيانات المتغيِّرين $X$ و$Y$ إيجابي وقوي.

الآن لو قمنا الآن بحساب مُعامل بيرسون للارتباط الخطي للبيانات الخام ذاتها (قبل صبها في جدول توزيع تكراري) سنجد باستخدام العلاقة [3,5-d] ما يلي:

$$\mathbf{r}_P=\frac{n\cdot\left(\sum_{i=1}^{n}x_i y_i\right)-\left(\sum_{i=1}^{n}x_i\right)\cdot\left(\sum_{i=1}^{n}y_i\right)}{\sqrt{n\cdot\left(\sum_{i=1}^{n}x_i^2\right)-\left(\sum_{i=1}^{n}x_i\right)^2}\cdot\sqrt{n\cdot\left(\sum_{i=1}^{n}y_i^2\right)-\left(\sum_{i=1}^{n}y_i\right)^2}}=0.987$$

وهذه العلاقة الأخيرة تُظهر لنا أنَّ الارتباط بين مجموعتي البيانات $X$ و$Y$ إيجابي وقوي جداً، فلو أمعنا النظر في تمثيل البيانات على لوحة الانتشار (انظر الشكل الآتي (١٢ ,٣)) فإنَّنا سنلاحظ قوة الارتباط بين المقرَّرين بسبب وقوع القسم الأكبر من البيانات على مستقيم واحد أو بالقرب منه. بالطبع إنَّ سبب وجود فارق بين القيمتين لـ $\mathbf{r}_P$ يكمن في عمليّة صبّ البيانات في جدول التوزيع التكراري حيث مُثِّلت البيانات المجمّعة بمراكز الفئات، وقد أشرنا إلى تأثير ذلك سابقاً في الفصل الثاني.




الشكل (١٢ ,٣)

٣- تَقدَّم ثمانية أشخاص لامتحان تحريري وآخر مقابلة في مقرَّر الإحصاء، حيث كانت نتائج الاختبار التحريري بالدرجات (الدرجة القصوى 60) في حين كانت نتائج اختبار المقابلة بالتقديرات (ممتاز- جيد جداً- جيد- وسط- ضعيف)، وكانت نتائجهم كما في الجدول الآتي:

الشكل (١٠ ,٣)

| الطالب | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| نتائج الاختبار التحريري $X$ | 48 | 38 | 32 | 50 | 50 | 38 | 55 | 55 |
| نتائج اختبار المقابلة $Y$ | جيد | وسط | وسط | جيد | جيد | وسط | جيد | جيد جداً |

ولنقم بدراسة الارتباط بين نتائج الامتحان التحريري والمقابلة لهؤلاء الأشخاص.

نلاحظ أنَّ البيانات المقدَّمة من النوع المختلط (كميّة ونوعيّة)، ولذلك سنستخدم مُعامل سبيرمان لارتباط الرتب، ومن أجل

ذلك سنضع رتباً بقيم متزايدة للبيانات معتمدين في ذلك الترتيب التنازلي (أو التصاعدي)، فمن أجل نتائج الاختبار التحريري سنبدأ بأول بيان (وهو القيمة 55) ونعطيه الرتبة 1 ثمَّ نعطي الرتبة 2للبيان المساوي له إن وجد وإلاَّ أعطيت للقيمة التي تليها، وهكذا حتى ننتهي من عمليّة إعط اء الرتب الأولية للبيانات الكميّة الناتجة عن الاختبار التحريري. وأمّا من أجل البيانات النوعيّة الممثّلة لاختبار المقابلة فإنَّنا نبدأ بأول بيان وهو جيد جداً (لكي يتماشى مع الأسلوب الذي استخدم من أجل نتائج الاختبار التحريري) فنعطيه الرتبة 1 ثمَّ نعطي الرتبة 2 للبيان المماثل له إن وجدت وإلاَّ أعطيت لأول جيد إن وجد، فإن لم تكن موجودة نعطيها لأول وسط، وهكذا حتى ننتهي من عملية إعطاء الرتب الأولية للبيانات النوعيّة الناتجة عن اختبار المقابلة، فنكون بذلك قد أنجزنا تعيين الرتب الأولية للبيانات. بعد ذلك نحسب المتوسط لرتب البيانات المتماثلة ونستخدمها كرتبة لكل من البيانات المتماثلة فنحصل على الرتب النهائية للبيانات.

الخطوة التالية نقوم بحساب الفروق بين الرتب النهائية المتقابلة ونقوم بحساب مربَّعاتها، فتكون نتائج هذه العمليّات كما في الجدول الآتي:

الشكل (٣, ١١)

| الطالب | A | B | C | D | E | F | G | H |
|---|---|---|---|---|---|---|---|---|
| نتائج الاختبار التحريري $X$ | 48 | 38 | 32 | 50 | 50 | 38 | 55 | 55 |
| الرتب الأولية $\alpha_i$ | 5 | 7 | 8 | 4 | 3 | 6 | 2 | 1 |
| الرتب النهائية $\alpha_i^*$ | 5 | 6.5 | 8 | 3.5 | 3.5 | 6.5 | 1.5 | 1.5 |
| نتائج اختبار المقابلة $Y$ | جيد | وسط | وسط | جيد | جيد | وسط | جيد | جيد جداً |
| الرتب الأولية $\beta_i$ | 2 | 6 | 7 | 3 | 4 | 8 | 5 | 1 |
| الرتب النهائية $\beta_i^*$ | 3.5 | 7 | 7 | 3.5 | 3.5 | 7 | 3.5 | 1 |
| $d_i = \alpha_i^* - \beta_i^*$ | 1.5 | -0.5 | 1 | 0 | 0 | -0.5 | -2 | 0.5 |
| $d_i^2$ | 2.25 | 0.25 | 1 | 0 | 0 | 0.25 | 4 | 0.25 |

فنجد أنَّ $\sum_{i=1}^{n} d_i^2 = 8$ ، ومن ثمَّ قيمة مُعامل سبيرمان لارتباط الرتب والخاص ببيانات $X$ و$Y$ تساوي:

$$\mathbf{r}_s = 1 - \frac{6 \cdot \sum_{i=1}^{n} d_i^2}{n \cdot (n^2 - 1)} = 1 - \frac{6 \times 8}{8(64-1)} = 0.905$$

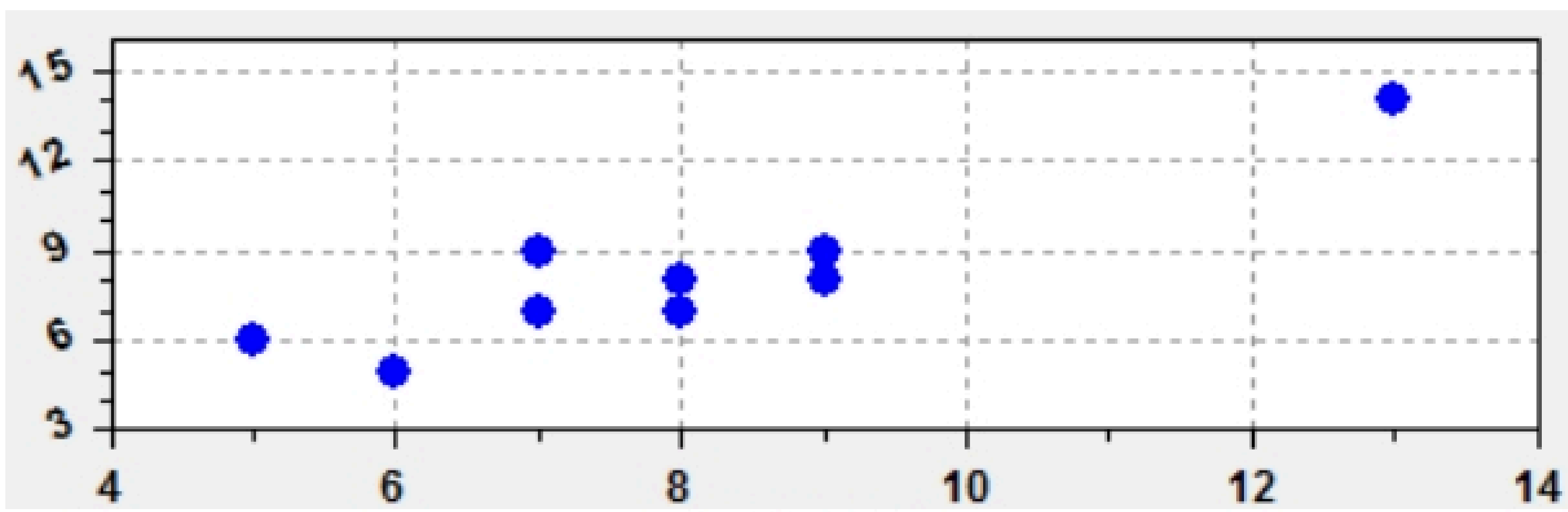
وهكذا نجد أنَّ الارتباط بين الاختبارين إيجابي وقوي جداً، وهذا يعني أنَّ نتائج الاختبارين التحريري والمقابلة لمقرَّر الإحصاء أظهرت أنَّ قدرات الأشخاص لا تتغيّر كثيراً بتغيّر نوع الاختبار.

٤- تمَّ تدريب عشرة أشخاص على نوع من التجارب، بحيث يدرَّب الشخص $k$ لعدد من المرّات يساوي $x_k$ مرَّة، وليكن $X$ متغيّرًا يرصد التكرارات التي قام بها المتدربون. بعد ذلك يَخضع كل متدرب لاختبار فيحصل على درجة $y_k$ من 15 (أي إن الدرجة القصوى 15)، ولنفترض أنَّ المتغيّر $Y$ يرصد التقييم بالدرجات للمتدربين. الآن بفرض أنَّ نتائج المتدربين كانت كما في الجدول الآتي:

الشكل (٣, ١٢)

| الشخص | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| عدد مرّات التدريب $X$ | 7 | 5 | 9 | 9 | 7 | 8 | 13 | 8 | 6 | 8 |
| درجة الاختبار $Y$ | 9 | 6 | 9 | 8 | 7 | 7 | 14 | 7 | 5 | 8 |

فنجد العرض الانتشاري للبيانات المقدَّمة في هذا الجدول السابق له الشكل الآتي.




الشكل (٣,١٣)

ولنقم بدراسة الارتباط الخطي بين المتغيِّرين $X$ و$Y$ باستخدام مُعامل سبيرمان الرتبي.

من أجل ذلك لنأخذ رتباً بقيم متزايدة للبيانات معتمدين في ذلك الترتيب التنازلي، كأن نبدأ من أجل قيم $X$ بأعلى قيمة وهي 13 فتُعطى الرتبة 1، ثمَّ نعطي الرتبة 2 للبيان 9 ومن ثمَّ 3 للبيان 9، ونتابع على هذا المنوال، فنحصل بذلك على الرتب الأوليّة للبيانات $X$، وبعد ذلك يُحسب المتوسط لرتب البيانات المتماثلة ويستخدم كرتبة لكل من هذه البيانات المتماثلة فنحصل على الرتب النهائية لبيانات $X$، وبشكل مماثل لما سبق نعالج مجموعة بيانات $Y$، فتصبح نتائج هذه العمليّات كما في الجدول الآتي:

الشكل (٣,١٣)

| الشخص | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| عدد مرَّات التدريب $X$ | 7 | 5 | 9 | 9 | 7 | 8 | 13 | 8 | 6 | 8 |
| الرتب الأولية $\alpha_i$ | 8 | 10 | 3 | 2 | 7 | 6 | 1 | 5 | 9 | 4 |
| الرتب النهائية $\alpha_i^*$ | 7.5 | 10 | 2.5 | 2.5 | 7.5 | 5 | 1 | 5 | 9 | 5 |
| درجة الاختبار $Y$ | 9 | 6 | 9 | 8 | 7 | 7 | 14 | 7 | 5 | 8 |
| الرتب الأولية $\beta_i$ | 3 | 9 | 2 | 5 | 8 | 7 | 1 | 6 | 10 | 4 |
| الرتب النهائية $\beta_i^*$ | 2.5 | 9 | 2.5 | 4.5 | 7 | 7 | 1 | 7 | 10 | 4.5 |
| $d_i = \alpha_i^* - \beta_i^*$ | 5 | 1 | 0 | -2 | 0.5 | -2 | 0 | -2 | -1 | 0.5 |
| $d_i^2$ | 25 | 1 | 0 | 4 | 0.25 | 4 | 0 | 4 | 1 | 0.25 |

فنجد أنَّ $\sum_{i=1}^{n} d_i^2 = 39.5$ ، وهكذا تكون قيمة مُعامل سبيرمان لارتباط الرتب للبيانات المقدَّمة تساوي:

$$\mathbf{r}_s = 1 - \frac{6 \cdot \sum_{i=1}^{n} d_i^2}{n \cdot (n^2 - 1)} = 1 - \frac{6 \times 39.5}{10(100-1)} = 0.761$$

وهكذا نلاحظ وجود ارتباط إيجابي قوي بين عدد مرَّات التدريب ودرجة الاختبار، وهذا يعني أنَّه كلما تلقى المتدرب دورات تدريبيّة أكثر فإنَّه يحصل على نتائج أفضل.

الآن، ومن أجل مقارنة معاملي بيرسون وسبيرمان للارتباط الخطي سنقوم بحساب مُعامل بيرسون للارتباط الخطي والخاص بالبيانات السابقة. حيث لدينا:

الشكل (٣,١٤)

| $i$ | $x_i$ | $x_i^2$ | $y_i$ | $y_i^2$ | $x_i \cdot y_i$ |
|---|---|---|---|---|---|
| 1 | 7 | 49 | 9 | 81 | 63 |
| 2 | 5 | 25 | 6 | 36 | 30 |
| 3 | 9 | 81 | 9 | 81 | 81 |
| 4 | 9 | 81 | 8 | 64 | 72 |
| 5 | 7 | 49 | 7 | 49 | 49 |
| 6 | 8 | 64 | 7 | 49 | 56 |
| 7 | 13 | 169 | 14 | 196 | 182 |
| 8 | 8 | 64 | 7 | 49 | 56 |
| 9 | 6 | 36 | 5 | 25 | 30 |
| 10 | 8 | 64 | 8 | 64 | 64 |
| sum | 80 | 682 | 80 | 694 | 683 |

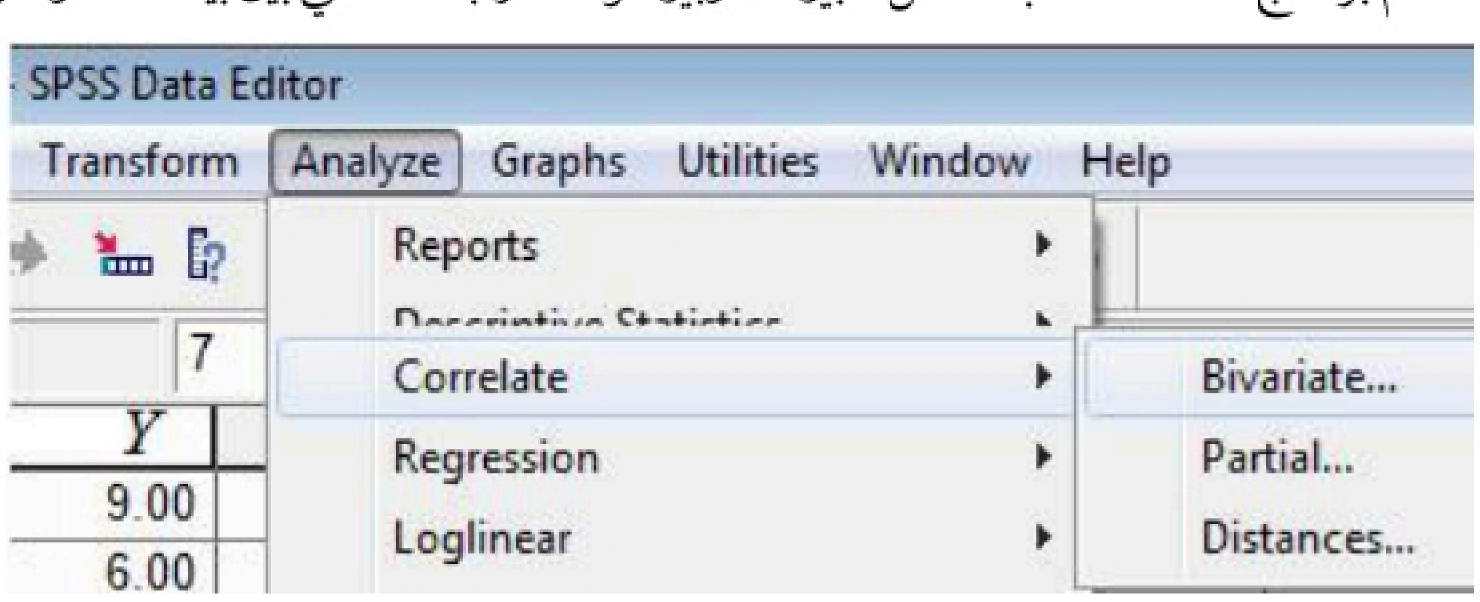
ومن ثمَّ نجد أنَّ قيمة مُعَامل بيرسون للارتباط الخطي بين بيانات $X$ و$Y$ تساوي:

$$\mathbf{r}_P = \frac{n\cdot\left(\sum_{i=1}^{n} x_i y_i\right)-\left(\sum_{i=1}^{n} x_i\right)\cdot\left(\sum_{i=1}^{n} y_i\right)}{\sqrt{n\cdot\left(\sum_{i=1}^{n} x_i^2\right)-\left(\sum_{i=1}^{n} x_i\right)^2}\cdot\sqrt{n\cdot\left(\sum_{i=1}^{n} y_i^2\right)-\left(\sum_{i=1}^{n} y_i\right)^2}} = \frac{430}{(20.494)(23.238)} = 0.903$$

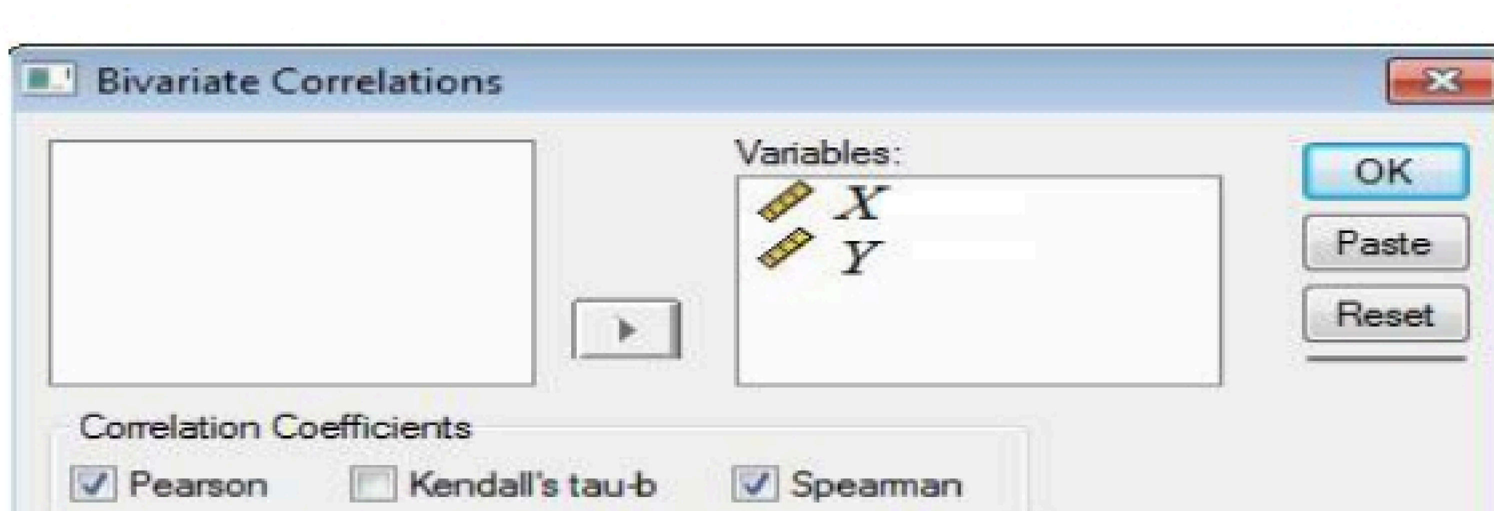
وهذا يعني وجود ارتباط إيجابي قوي جداً بين عدد مرَّات التدريب ودرجة الاختبار.

لاحظ أنَّ الفارق بين نتيجتي المُعاملين (لـ بيرسون وسبيرمان) هو أمر طبيعي ومتوقع أيضاً لأنَّ مُعامل سبيرمان لا يأخذ قيم البيانات بالحسبان وإنّما يعتمد على رتبها فقط.

كما يمكن استخدام برنامج SPSS لحساب مُعامل سبيرمان وبيرسون للارتباط الخطي بين بيانات $X$ و$Y$ وفقاً للخطوات الآتية:




ومن ثمَّ



فنحصل على تقرير فيه الجزء الذي يهمّنا هو الآتي:

| | | | $X$ | $Y$ |
|---|---|---|---|---|
| Spearman's rho | $X$ | Correlation Coefficient | 1.000 | .752 |

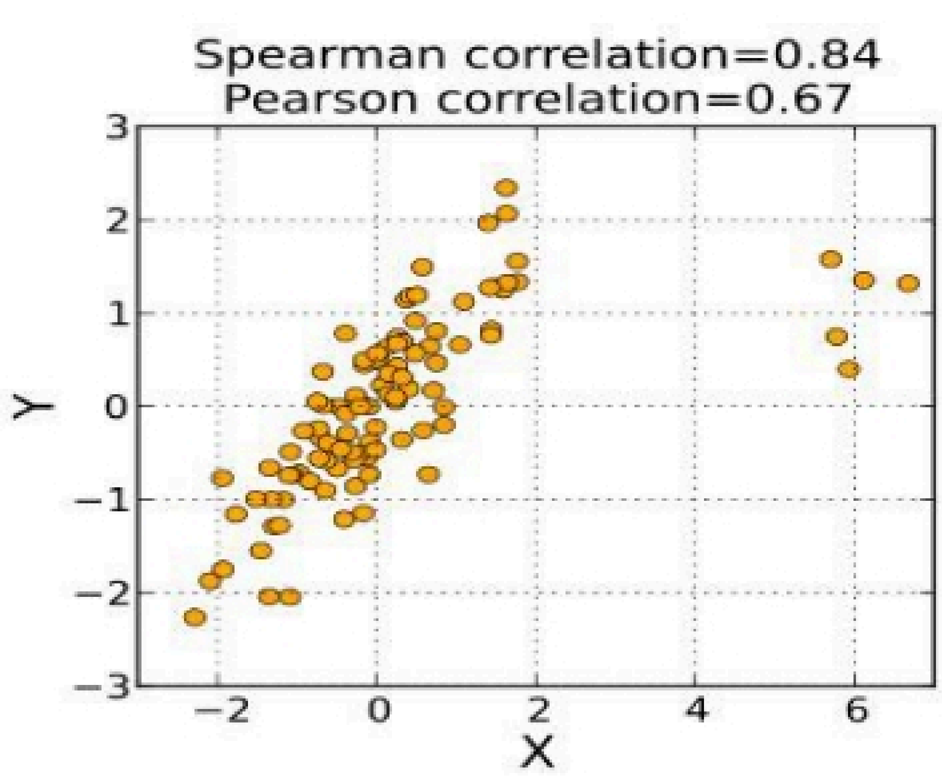
الذي يعطينا $\mathbf{r}_s = 0.752$، وهذه القيمة تختلف قليلاً عن القيمة التي قمنا بحسابها لهذا المُعامل، وسبب هذا الفرق يعود لاستخدام البرنامج صيغة معدَّلة في عمليّة الحساب.

### (٣,٥,١,٣) ملاحظات

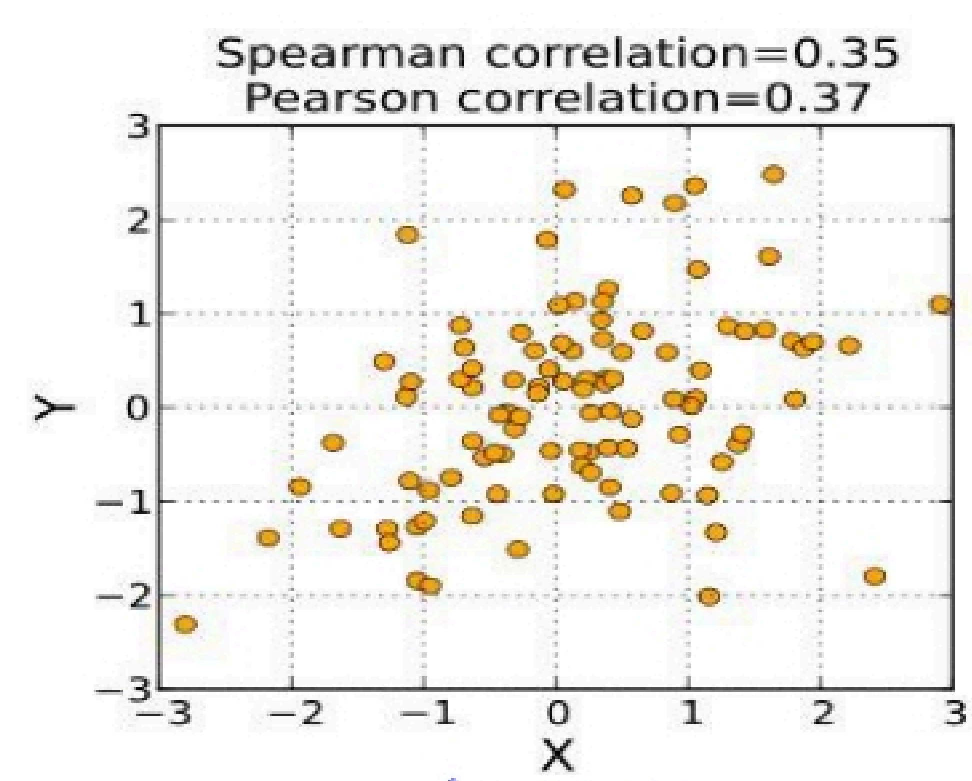
١- يُعدّ مُعامل بيرسون للارتباط الخطي أكثر دقّة من مُعامل سبيرمان لارتبط الرتب في تعيين درجة الارتباط الخطي، بمعنى أنّه إذا كانت البيانات التي هي قيد الدراسة هي بيانات كميّة كاملة (أي لا يوجد فيها بيانات مفقودة)، فعندئذ من المفضّل استخدام مُعامل بيرسون للارتباط الخطي.

٢- إذا لاحظنا وجود بيانات مفقودة ولكن ترتبها بين البيانات وتبعيتها للعناصر المدروسة معلوماً، فعندئذ يمكن اللجوء إلى استخدام مُعامل سبيرمان لارتباط الرتب.

٣- إذا كانت مجموعة البيانات لا تحتوي على بيانات متطرّفة (أو شاذّة في سلوكها) فعندئذ ستكون قيمتي مُعامل بيرسون للارتباط الخطي ومُعامل سبيرمان لارتباط الرتب قريبتين من بعضهما (انظر الشكل (١٢,٣ أ)).




الشكل (١٤,٣ ب)




الشكل (١٤,٣ أ)

وأمّا إذا كانت مجموعة البيانات تحتوي على بعض البيانات المتطرّفة (أو الشاذّة في سلوكها)، فعندئذ سيكون هناك فارق (قد يكون كبيراً نسبياً) بين قيمتي مُعامل بيرسون للارتباط الخطي ومعامل سبيرمان لارتباط الرتب (انظر الشكل السّابق (١٢,٣ ب))، ولذلك يفضّل استخدام مُعامل سبيرمان لارتباط الرتب في هذه الحالة بسبب تأثر المتوسط الحسابي (ومن ثمَّ مُعامل بيرسون للارتباط الخطي) بالقيم المتطرّفة.

٤- إن مُربَّع قيمة مُعامل الارتباط (سواء كانت لعيِّنة أو لمجتمع) تُدعى مُعامل التحديد Coefficient of Determination، ويُرمز له بـ $\mathbf{r}^2$ أو $\rho^2$ (وفي بعض البرامج الإحصائية يُشار إليه بـ **R-square**)، وهذه القيمة تستعمل كتقدير لقوة علاقة الارتباط بين متغيِّرين أيضاً. إنَّ النسبة المئوية لمعامل التحديد $\mathbf{r}^2\,\%$ تشكِّل مقدار تَغيُّر الظاهرة $X$ المرتبطة مع تَغيُّر $Y$، فعلى سبيل المثال (في الكيمياء الحيوية) إذا كانت قيمة مُعامل الارتباط لكميّة الكلوروفيل $\alpha$ والمادة العضوية تساوي $\mathbf{r}$=0.50 فإنَّ مُعامل التحديد $\mathbf{r}^2 = 0.25$ يدلّ على أنَّ 25% من التغيُّر في أحد المتغيِّرين مرتبطة بالمتغيِّر الآخر (أو تُوضَّح من خلال تغيُّر المتغيِّر الآخر)، أو بتعبير آخر نقول إنَّ 75% من التَغيُّر في أحد المتغيِّرين لم توضَّح بوساطة التغيّر الذي حصل للمتغيِّر الآخر.

### (٣.١,٦) مُعامل الاقتران Coefficient of Association

في كثير من الأحيان (وخاصّة لدى الدراسات الاجتماعيّة أو السلوكيّة) تكون البيانات المُعطاة مُجمّعةً في جدول ولكنَّ هذه البيانات نوعيّة مثل: قوي وضعيف، ذكي وغبي، ذكر وأنثى وعلى هذا المنوال. بالطبع قد قمنا سابقاً بتقديم مُعامل سبيرمان لارتباط الرتب الذي يمكن استخدامه من أجل البيانات النوعيّة التي يمكن ترتيبها، وأمّا من أجل الحالة التي لا يمكن ترتيب البيانات فيها، أو إن رُتبت فلا أهميّة لترتيبها، فإن مُعامل سبيرمان لارتباط الرتب يُصبح غير قابل للتطبيق، ولذلك لابدّ من إيجاد مقاييس خاصّة تتعامل مع هـذه البيـانات.

لقد وضع الإحصائي البريطاني يول (1871 –1951) George Udny Yule مقياساً لدراسة درجة ونوعيّة الارتباط بين ظاهرتين بياناتها نوعيّة أو كميّة وتنقسم كل منها إلى نوعين أو مجموعتين فقط (مثل: قوي وضعيف – سريع وبطيء و...)، وأسماه مُعَامل الاقتران، وفي مثل هذه الحالة تكون البيانات المتاحة للباحث كما في الجدول الآتي:

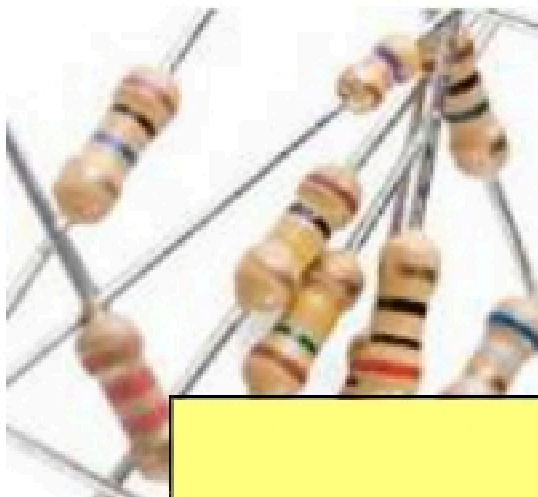
الشكل (٣,١٥)

| | | الظاهرة الأولى $X$ | |
|---|---|---|---|
| | | عدد العناصر التي حَقّقت الظاهرة الأولى | عدد العناصر التي لم تُحقّق الظاهرة الأولى |
| الظاهرة الثانية $Y$ | عدد العناصر التي حَقّقت الظاهرة الثانيّة | $a$ | $b$ |
| | عدد العناصر التي لم تُحقّق الظاهرة الثانيّة | $c$ | $d$ |

عندئذ يُحسب مُعَامل الاقتران بين هاتين الظاهرتين من خلال العلاقة الآتية:

$$\mathbf{r}_{CA} := \frac{ad - bc}{ad + bc} \qquad [3,10]$$

علماً أنَّ اقتراب قيمة $\mathbf{r}_{CA}$ من القيمة 0 تعني ضعف الاقتران بين الظاهرتين، وأمّا اقتراب قيمة $\mathbf{r}_{CA}$ من القيمة $\pm 1$ فإنّها تعني اقتراب درجة الاقتران من الكمال بين الظاهرتين إيجابيًّا (عندما تكون الإشارة موجبة) أو عكسيًّا (عندما تكون الإشارة سالبة).



#### (٣,١,٦,١) مثال

في دراسة لمعرفة مدى تأثير مرور تيار أعلى من المسموح به في مقاومة أوميّة ذات استطاعة 0.5W واط توفرت لدينا البيانات الآتية:

الشكل (٣,١٦)

| | | الظاهرة الأولى (العريض للتيّار) | |
|---|---|---|---|
| | | عدد العناصر التي عُرِّضت لتيار أعلى من المسموح به | عدد العناصر التي لم تُعرَّض لتيار أعلى من المسموح به |
| الظاهرة الثانية (الإصابة بالتلف) | عدد العناصر التي أصيبت بالتلف | **131** | **78** |
| | عدد العناصر التي لم تصب بالتلف | **52** | **318** |

فنجد من هذه البيانات أنَّ:

$$\mathbf{r}_{CA} = \frac{131\times318-78\times52}{131\times318+78\times52} = \frac{37602}{45714} = 0.823$$

وهذا يعني وجود ارتباط طرديّ قويّ بين تَعرّض المقاومة لتيّار أعلى من المسموح به وتلفها، أو يُقال إنَّ تلف المقاومة غير مستقل عن تعرضّها لتيّار أعلى من المسموح به.

قد يكون لدينا ظواهر ذات بيانات نوعيّة أو كميّة، وتتجزأ في أكثر من نوعين أو مجموعتين، فعندئذ نلاحظ أنَّ مُعامل الاقتران لا يقدِّم لنا أية مساعدة في هذا المجال، ولذلك لا بدّ من مقياس آخر يتعامل مع هذه المُعطيات. لقد قام بيرسون بوضع مقياس لدراسة درجة ونوعية الارتباط لمثل هذه المُعطيات، وأُطلق عليه اسم مُعامل التوافق الذي تقدِّمه الفقرة الآتية.

### (٣,١,٧) مُعامل التوافق Coefficient of Contingency

لتكن لدينا مجموعة بيانات حجمها $n$ ومُعطاة كما في الجدول الآتي:

الشكل (٣,١٧)

| | | بيانات المتغيِّر $X$ | | | | Total |
|---|---|---|---|---|---|---|
| | | $x_1$ | $x_2$ | ..... | $x_k$ | |
| بيانات المتغيِّر $Y$ | $y_1$ | $a_{11}$ | $a_{12}$ | ..... | $a_{1k}$ | $n_{1\bullet}$ |
| | $y_2$ | $a_{21}$ | $a_{22}$ | ..... | $a_{2k}$ | $n_{2\bullet}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | ..... | $\vdots$ | $\vdots$ |
| | $y_\ell$ | $a_{\ell 1}$ | $a_{\ell 2}$ | ..... | $a_{\ell k}$ | $n_{\ell\bullet}$ |
| Total | | $n_{\bullet 1}$ | $n_{\bullet 2}$ | ..... | $n_{\bullet k}$ | $n$ |

علماً أنَّ $a_{ij}$ هي القيمة الواقعة في العمود $i$ والصف $j$ وتشير إلى عدد البيانات الموافقة لـــ $x_i$ و $y_j$ معاً، وأمّا $n_{j\bullet}$ و $n_{\bullet i}$ فإنَّهما يحسبان من أجل كل $i \in \mathbb{N}_k$ و $j \in \mathbb{N}_\ell$ كما يلي:

$$n_{j\bullet} = \sum_{i=1}^{k} a_{ji} \qquad \& \qquad n_{\bullet i} = \sum_{j=1}^{\ell} a_{ji}$$

عندئذ يُعطى مُعامل التوافق لهذه البيانات (وسنرمز له بـ $\mathbf{r}_{CC}$) بالعلاقة الآتية:

$$\mathbf{r}_{CC} = \sqrt{\frac{G}{G+n}} \qquad [3,11]$$

علماً أنَّ $G$ تُحسب كما يلي:

$$G = n\left(\sum_{i=1}^{k}\sum_{j=1}^{\ell}\frac{(a_{ij})^2}{n_{j\bullet}\cdot n_{\bullet i}}\right) - n \qquad [3,12]$$

والمثال الآتي يوضِّح ذلك.

#### (٣,١,٧,١) مثال

إنَّ الجدول الآتي يَعرض لنا المهارات المخبرية لعيِّنة من 473 طالباً في تنفيذ تجربتين $X$ و$Y$.

الشكل (٣,١٨)

| | | بيانات التجربة $X$ | | | | |
|---|---|---|---|---|---|---|
| | | ممتاز | جيد جداً | جيد | مقبول | Total |
| بيانات التجربة $Y$ | مقبول | 14 | 23 | 18 | 10 | 65 |
| | جيد | 50 | 82 | 103 | 45 | 280 |
| | جيد جداً | 21 | 30 | 50 | 27 | 128 |
| Total | | 85 | 135 | 171 | 82 | 473 |

ولنبيِّن نوعيّة ودرجة الارتباط بين هاتين التجربتين.

من أجل ذلك لنستخدم صيغة مُعَامل التوافق (العلاقة [3,11]) فنجد أنَّ:

$$G = n\left(\sum_{i=1}^{k}\sum_{j=1}^{\ell}\frac{\left(a_{ij}\right)^2}{n_{j\bullet}\cdot n_{\bullet i}}\right)-n = 473\left(\frac{(10)^2}{(82)\cdot(65)}+\frac{(18)^2}{(171)\cdot(65)}+\frac{(23)^2}{(135)\cdot(65)}+\frac{(14)^2}{(85)\cdot(65)}\right.$$

$$\left.+\frac{(45)^2}{(82)\cdot(280)}+\dots+\frac{(30)^2}{(135)\cdot(128)}+\frac{(21)^2}{(85)\cdot(128)}\right)-473$$

$$= 473\times 1.0066-473 = 3.1218$$

ومن ثمَّ نجد قيمة مُعامل التوافق لهذه البيانات هي:

$$\mathbf{r}_{CC}=\sqrt{\frac{G}{G+n}}=\sqrt{\frac{3.1218}{3.1218+473}}=0.081$$

وهذا يعني أنَّ التجربتين ترتبطان مع بعضهما ارتباطاً إيجابياً، ولكنَّ درجة هذا الارتباط ضعيفة جداً.

## (٣,٢) الانحدار الخطي البسيط

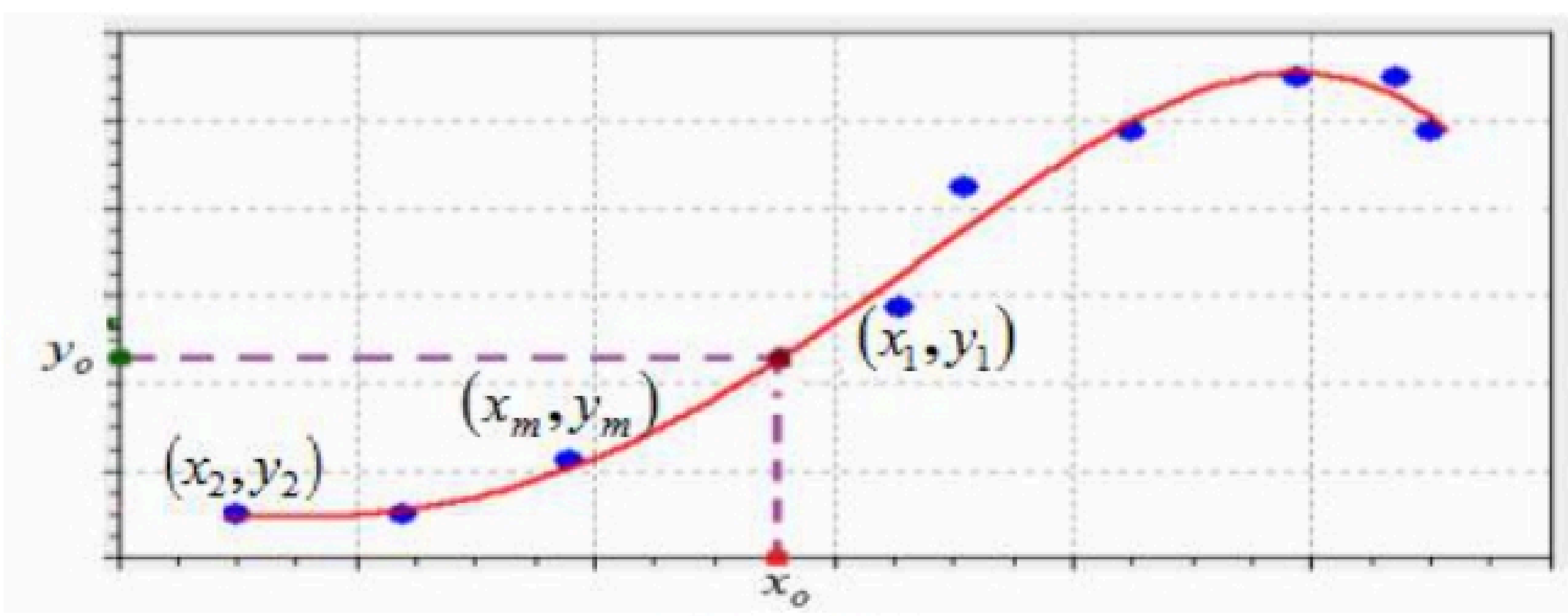
### Simple Linear Regression

قمنا فيما سبق بدراسة الارتباط لظاهرتين ومن ثمَّ قياس درجة الارتباط وفق طرائق مختلفة تتناسب كل منها وحالة معيَّنة لطبيعة البيانات، ولكن إذا توقفنا عند هذه المرحلة فإنَّ الفائدة ستكون محدودةً، وذلك أنَّه من الأفضل (وبناءً على وجود تلك العلاقة للارتباط بين ظاهرتين) استخدام تلك الحقيقة في تقدير قيمة متغيِّر إذا ما علمنا قيمة أو أكثر للمتغيِّر الآخر. إنَّ تقدير قيمة متغيِّر إذا ما عُلمت علاقة الارتباط بين هذا المتغيِّر ومتغيِّر آخر مبنيٌّ على ما يُسمَّى بتحليل الانحدار. في الواقع إنَّ الانحدار هو تمثيل للعلاقة المتوسطة بين المتغيِّرات، فإذا وجدت علاقة تربط بين المتغيِّرات المطلوب دراستها فإنَّه يمكن توفيق معادلة منحنى أو خط يُحدِّد طبيعة تلك العلاقة، ومن ثمَّ يُمكننا استخدام ذلك المنحنى أو الخط لتقدير القيم النظرية لمتغيِّر إذا عُلمت القيم للمتغيِّر الآخر. هكذا نجد أنَّه يمكن استخدام تحليل الانحدار لتقدير قيمة متغيِّر تابع (المتغيِّر الذي نودّ تعليله) بناءً على قيمة واحدة على الأقل من قيم المتغيِّر المستقل (المتغيِّر المستخدم لتعليل المتغيِّر التابع)، ومن ثمَّ إمكانية شرح تأثير التغيّرات في المتغيِّر المستقل على المتغيِّر التابع.

أمَّا من حيث نشأة مفهوم الانحدار، فقد جاءت من دراسات في علم الوراثة للإحصائي وعالم الوراثة والنفس والاجتماع الإنكليزي غالتون (Sir Francis Galton (1822-1911) (خلال أواخر القرن التاسع عشر) عند تقديم دراسته حول العلاقة التي تربط بين أعمار الآباء والأبناء لأجيال متعاقبة عديدة، وذلك بغية استقراء واقع تطور العمر لدى العنصر البشري، فوجدَ أنَّ أعمار الأبناء في حالة انحسار متواصل عبر الأجيال المتعاقبة، وكذلك لاحظ غالتون أن خصائص مُحدَّدة مثل الطول في الآباء لا تنتقل تماماً إلى ذرياتهم (الطول في حالة انحسار أيضاً)، وعُبِّر عن ذلك بالقول: إنَّ أعمار الأبناء في حالة انحدار لأنَّ التمثيل البياني لها كان منحدراً، ومن هنا جاءت تسـمية

تحليل الانحدار. في الواقع من الممكن أنَّ تكون العلاقة التي تربط بين متغيِّرات أخرى متعلِّقة بعناصر عيِّنة (أو مجتمع) ما هي علاقة لها مَنْحى صاعد وليس منحدرًا (أوهابطًا) كالعلاقة التي تربط بين الطول والوزن لمجموعة من البشر.

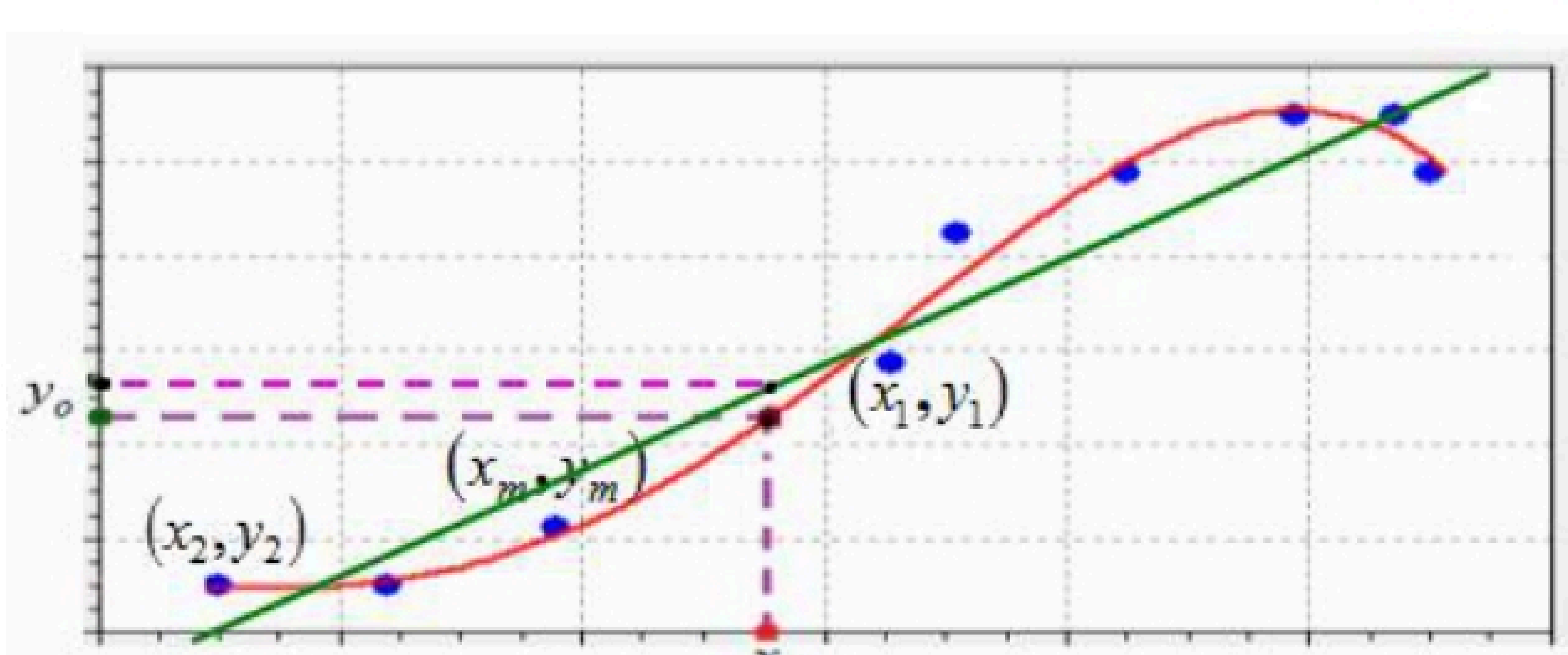
لنفترض الآن أنَّه لدينا $(x_1,y_1)$، $(x_2,y_2)$، ...و $(x_n,y_n)$ مجموعة منتهية من المشاهدات الناتجة عن متغيِّرين $X$ و$Y$، ولنفترض على سبيل المثال أنَّ تمثيلها على لوحة الانتشار له العرض المقدَّم في الشكل الآتي (١٥ ,٣ .أ)، فعندئذ من أولى مهام تحليل الانحدار تَمكيننا من تقدير (أو تخمين) قيمة $y_o$ مقابلة لقيمة $x_o$ مجهولة تقع ضمن نطاق البيانات الأصلية $x_1$، $x_2$، ... و $x_n$ ، وكذلك يساعدنا في بعض الحالات على التنبؤ بقيمة $y_o$ مقابلة لقيمة $x_o$ التي تقع خارج نطاق البيانات الأصلية $x_1$، $x_2$، ... و $x_n$ .



الشكل (١٥ ,٣ .أ)

بالطبع إنَّ أفضل تقدير لهذه القيمة ينتج عن استخدام الدالَّة التي يمرّ رسمها البياني من جميع النقاط الممثِّلة للبيانات، ولكن تعيين معادلة هذه الدالَّة في الحالة العامّة قد يكون صعباً جداً إن لم يكن غير ممكن في بعض الحالات، ولذلك يبحث المرء في تعيين الدالَّة التي رسمها البياني يمرّ من بعض هذه النقاط وبالقرب من بعضها الآخر إن أمكن ذلك (كما في الشكل (١٥ ,٣ .أ))، وتعيين مثل هذه الدالَّة يحتاج في معظم الأحيان إلى برامج خاصّة بتوفيق المنحنيات Curves Fitting.

سنبحث فيما يلي في أبسط أنواع العلاقات التي تربط بين متغيِّرين $X$ و$Y$، وهي تُدعى **معادلة الانحدار الخطي البسيط**، وعبارة بسيط هنا تعني حالة وجود متغيِّرين فقط. أي إنَّنا سنسعى لتعيين الدالَّة التي رسمها البياني مستقيم يمرّ من جميع هذه النقاط أو يمرّ من بعض هذه النقاط وبالقرب من بعضها الآخر إن أمكن ذلك، وإن لم يكن ذلك ممكناً فليكن المستقيم الذي يمرّ بالقرب من هذه النقاط كما في الشكل (١٥ ,٣ .ب) الآتي:



الشكل (١٥ ,٣ .ب)

في الحقيقة توجد طرائق عديدة لتعيين العـلاقة الخطّية التي تربط بين متغيِّرين $X$ و$Y$ وسنتناول هنا إحدى هذه الطرائق التي تُدعى طريقة المربَّعات الصغرى.

### (١, ٢, ٣) طريقة المربَّعات الصغرى Least Square Method

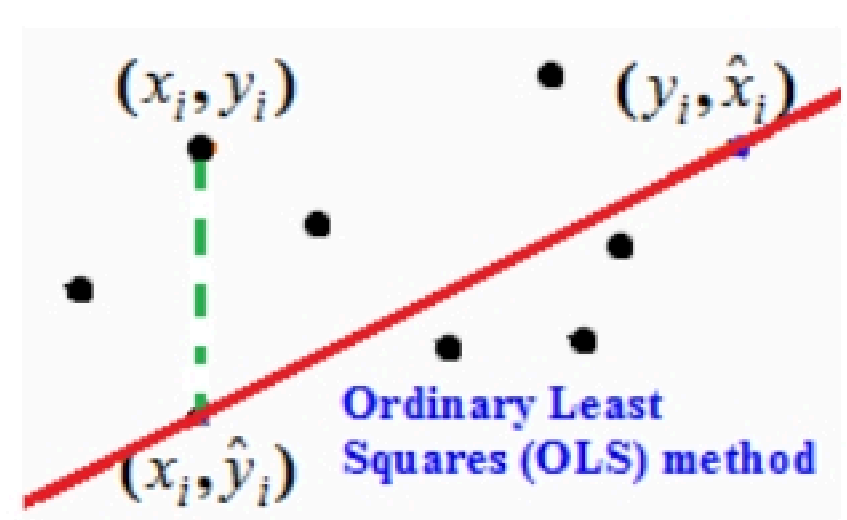
لقد نُشرت أول مقالة واضحة وموجزة لطريقة المربعات الصغرى من قبل الرياضياتي الفرنسي ليجاندر Adrien-Marie Legendre (1752-1833) في عام 1805، حيث وصفت هذه التقنية على أنَّها الإجراءات الجبرية لتوفيق معادلات خطيّة للبيانات، وقد استعرض ليجاندر طريقته الجديدة من خلال تحليل البيانات نفسها كما فعل لابلاس لشكل الأرض. لقد لاقت طريقة المربعات الصغرى المقدّمة من ليجاندر اعترافاً مباشراً من قبل كبار علماء الفلك والمساحة (الجيوديسيا) في ذلك الزمان. إنَّ طريقة المربعات الصغرى هي إحدى الطرائق المستخدمة في توفيق المنحنيات لمجموعة منتهية من المشاهدات المُعطاة (على شكل ثنائيات) وناتجة عن تطبيق دراسة متغيِّرين $X$ و$Y$ على مجموعة من العناصر، وتقوم هذه الطريقة على ركيزتين هما:

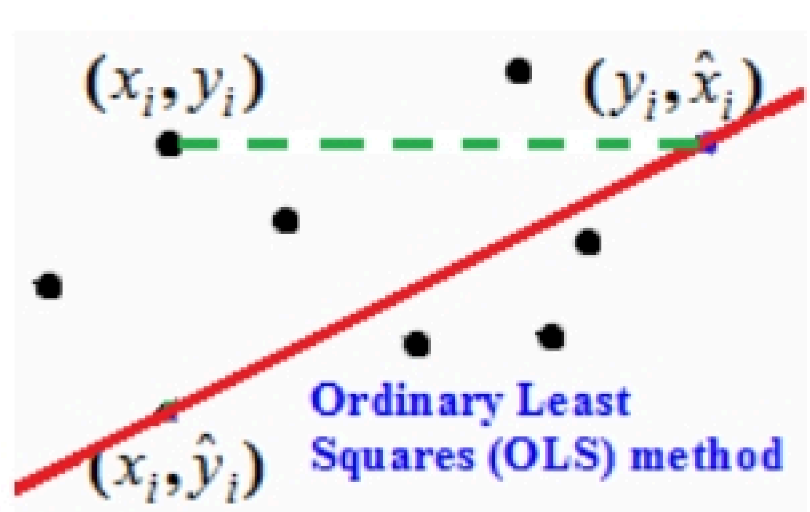
١- مجموع انحرافات المشاهدات عن النقاط المقابلة لها من هذا المستقيم يساوي الصفر دوماً.

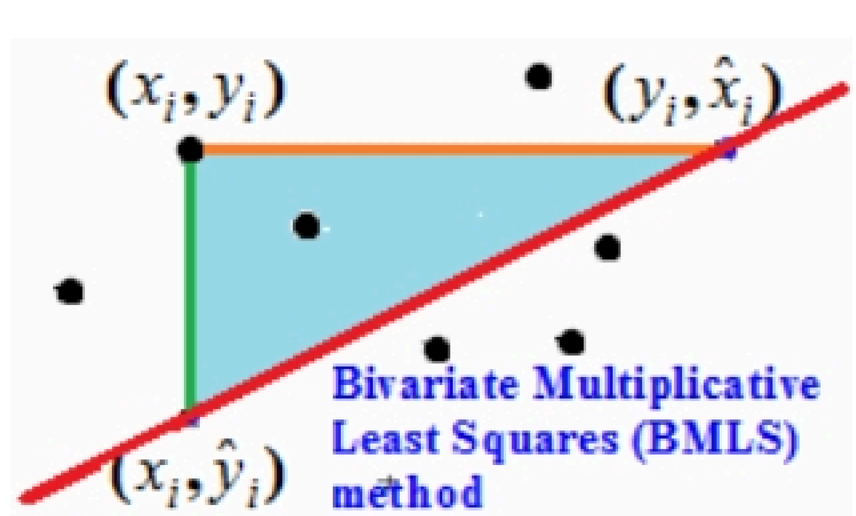
ب- جعل مجموع مربّعات انحرافات قيم المشاهدات عن المستقيم المُوَفِّق للمشاهدات أصغر ما يمكن.

عندئذ ينظر إلى هذا المستقيم على أنّه أفضل مستقيم يمرّ بالنقاط الممثِّلة للمشاهدات وفقًا لطريقة المربّعات الصغرى. بالطبع توجد طرائق عديدة لتفسير البندين (أ) و(ب) السّابقين (انظر العروض الآتية لتوضيح ذلك) ومنها:

١- طريقة **OLS**: تقوم هذه الطريقة على جعل $\sum_i (y_i - \hat{y}_i)^2$ أو $\sum_i (x_i - \hat{x}_i)^2$ أصغر ما يمكن. علماً أنَّ $\hat{y}_i$ هي القيمة المُقدَّرة لـ $y$ عند إعطاء قيمة $x_i$، وأمّا $\hat{x}_i$ فهي القيمة المُقدَّرة لـ $x$ عند إعطاء قيمة $y_i$ وفقاً لمعادلة المستقيم المُوَفِّق للمشاهدات.

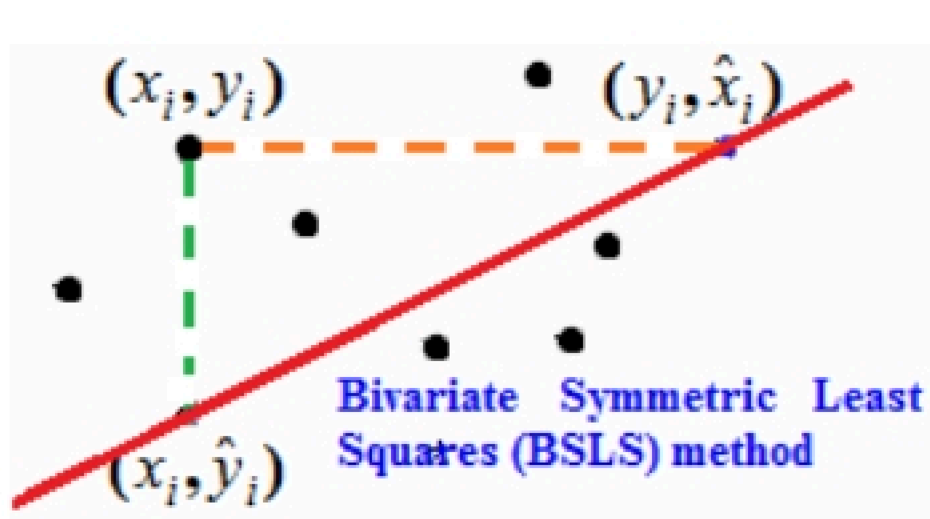


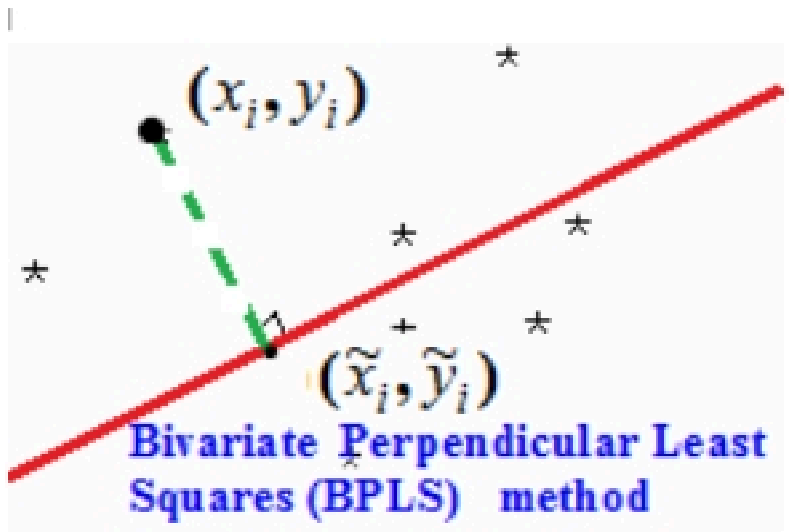

٢- طريقة **BMLS**: تقوم هذه الطريقة على جعل المقدار الآتي أصغر ما يمكن:

$$\sum_i \left| (y_i - \hat{y}_i) \cdot (x_i - \hat{x}_i) \right|$$




٣- طريقة **BSLS**: تقوم هذه الطريقة على جعل المجموع أصغر ما يمكن:

$$\sum_i \left[ (y_i - \hat{y}_i)^2 + (x_i - \hat{x}_i)^2 \right]$$




٤- طريقة **BPLS**: تقوم هذه الطريقة على جعل مجموع مربّع المسافات الناتجة عن الإسقاط العمودي للنقاط $(x_i, y_i)$ على مستقيم المُوَفّق للمشاهدات أصغر ما يمكن.

فيما يلي سنعتمد على طريقة المربعات الصغرى العادية Ordinary Least Squares Method (طريقة OLS) فقط لتعيين معادلة مستقيم الانحدار، وسنذكرها باسم **طريقة المربّعات الصغرى** على سبيل التبسيط، ووفقاً لهذه الطريقة سنناقش الحالتين الموافقتين لبيانات عيِّنة ومن ثمَّ لبيانات مجتمع.

## (٢ , ٢ , ٣) معادلتا مستقيما الانحدار لبيانات عيِّنة Equations of Straights Regression for Sample Data

بفرض أنَّ $(x_1, y_1)$، $(x_2, y_2)$،...،و $(x_n, y_n)$ بيانات عيِّنة أخضعت لدراسة متغيِّرين $X$ و$Y$، فعندئذ لتعيين مستقيم المربّعات الصغرى لمجموعة البيانات المُعطاة سنميّز بين نوعين من المستقيمات هما:

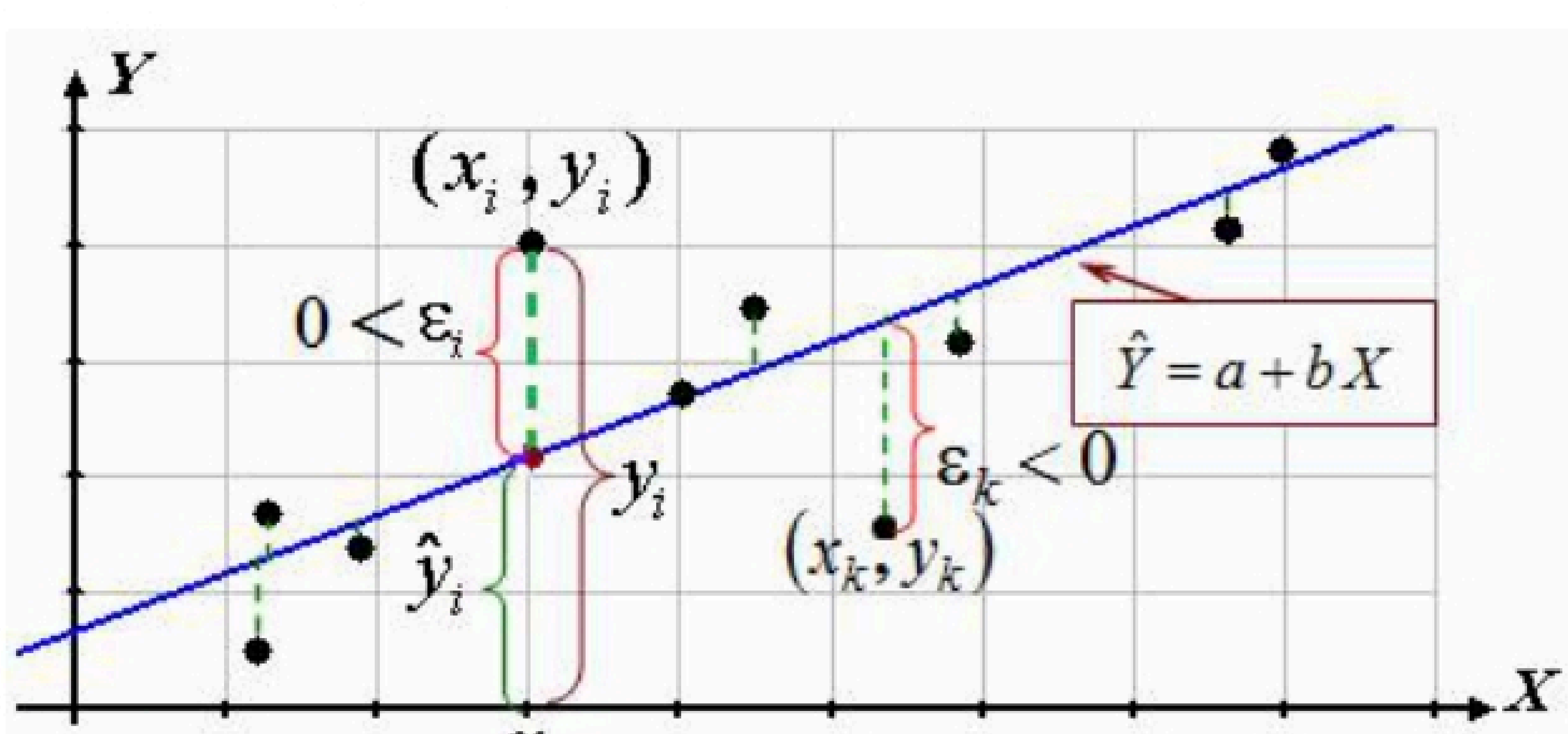
### (١ , ٢ , ٢ , ٣) معادلة مستقيم انحدار $Y$ على $X$ Equation of Straight Regression of $Y$ on $X$.

إذا كان الهدف تعيين قيمة $\hat{y}_i$ كتقدير لقيمة $y_i$ باستخدام هذا المستقيم لدى إعطاء قيمة محدّدة لـ $x_i$ (انظر الشكل الآتي)، فعندئذ يكون لمعادلة مستقيم المربعات الصغرى لهذه البيانات العرض الآتي:

$$\hat{Y} = a + b\,X \qquad [3,13\text{-}a]$$

ويُدعى المستقيم النّاتج عن هذه العلاقة بـ **مستقيم انحدار $Y$ على $X$**، وهي تساعدنا في تقدير قيم $Y$ لدى إعطاء قيم لـ $X$، بمعنى أنَّه من أجل $x_i$ قيمة مُحدَّدة لـ $X$ فإنَّ القيمة $\hat{y}_i$ المُقدَّرة لـ $y_i$ تحسب وفقاً للصيغة الآتية:

$$\hat{y}_i = a + b\,x_i \qquad ; i \in \mathbb{N}_n \qquad [3,13\text{-}b]$$




الشكل (٣ , ١٦)

إنَّ الفارق بين القيمة الحقيقية $y_i$ والقيمة المُقدَّرة لها $\hat{y}_i$ وفقاً لمعادلة المستقيم [3–13-a] يُدعى **الباقي** Residual أو **الخطأ** Error المرافق لتقدير قيمة $y_i$ بالقيمة $\hat{y}_i$ عند استخدام معادلة المستقيم المُعطى بالعلاقة [3–13-a]، ويُرمز له بـ $\varepsilon_i$، ومن ثمَّ يمكننا كتابة الصيغة التي تُعطينا البواقي على النحو الآتي:

$$\varepsilon_i = y_i - \hat{y}_i \qquad ; i \in \mathbb{N}_n$$

وهذا المقدار يمكن أن يكون موجباً أو سالباً أو صفراً وذلك بحسب موضع النقطة من المستقيم أهي أعلى المستقيم أو أسفله أم واقعة عليه (انظر الشكل السابق (٣ , ١٦)).

إنَّ قيم متغيِّر الاستجابة $\hat{Y}$ الذي يلعب دور المتغيِّر التابع في المعادلة [3–13-b] تُدعى القيم المقدَّرة (أو المتنبأ بها) Predicted Values (أو القيم التي تمَّ توفيقها Fitted Values)، وأمّا المتغيِّر التفسيري $X$ الذي له دور المتغيِّر المستقل في معادلة المستقيم فإنَّه يُدعى المتغيِّر المُتنبأ به Predictor Variable. في حين أنَّ $a$ و$b$ هما ثابتان حقيقيان يُطلب تعيينها ويُدعيان مُعَاملي الانحدار (أو وسطي الانحدار) Regression Coefficients، علماً أنَّ $a$ يمثِّل قيمة تقاطع مستقيم الانحدار مع المحور $\mathbf{oY}$ في حين يمثِّل $b$ قيمة الميل لهذا المستقيم.

الآن (وكما ذكرنا سابقاً) لتعيين قيم $a$ و $b$ في المعادلة [3–13-a] وفقاً لطريقة المربعات الصغرى يجب أن يكون ما يلي محقَّقاً:

$$\textbf{i)}\ \sum_{i=1}^{n} \varepsilon_i = 0 \qquad \& \qquad \textbf{ii)}\ \sum_{i=1}^{n} \varepsilon_i^2 = \textbf{minimal}$$

عندئذ من تحقُّق الشرط $\sum_{i=1}^{n} \varepsilon_i = 0$ ينتج لدينا ما يلي:

$$\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right) = 0 \;\Rightarrow\; \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} \hat{y}_i = b\sum_{i=1}^{n} x_i + n\,a \qquad \mathbf{(1)}$$

ولكي يكون الشرط (ii) مُحقَّقاً يلزم أنَّ يكون (وذلك باستخدام الاشتقاق الجزئي):

$$\frac{\partial}{\partial b}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 = 0 \;\Rightarrow -2\sum_{i=1}^{n} x_i\left(y_i - (bx_i + a)\right) = 0$$

ومنها نجد:

$$\sum_{i=1}^{n} x_i\, y_i = b\sum_{i=1}^{n} x_i^2 + a\sum x_i \qquad \mathbf{(2)}$$

وكذلك لدينا:

$$\frac{\partial}{\partial a}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2 = 0 \quad\Rightarrow\quad -2\sum_{i=1}^{n}\left(y_i - (bx_i + a)\right) = 0 \quad\Rightarrow\quad \sum_{i=1}^{n} y_i = b\sum_{i=1}^{n} x_i + n\,a$$

وهذه الأخيرة تُطابق العلاقة (1)، ومنه بحل جملة المعادلتين (1) و (2) بالنسبة إلى $a$ و$b$ نجد أنَّ:

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad [3,14\text{-}a]$$

وهذه يُمكن كتابتها بدلالة قيمة مُعَامل الارتباط الخطي $\mathbf{r}$ والانحراف المعياري لكل من $X$ و$Y$ على النحو الآتي أيضاً:

$$b = \frac{S_{X,Y}}{S_X^2} = \frac{S_Y}{S_X}\,\mathbf{r} \qquad [3,14\text{-}b]$$

وهذه القيمة التي تُمثِّل ميل مستقيم الانحدار تُدعى مُعَامل انحدار $Y$ على $X$، وأمّا قيمة $a$ التي تمثِّل قيمة تقاطع المستقيم مع المحور $\mathbf{oY}$ فيكون لها العرض الآتي:

$$a = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} \qquad [3,15\text{-}a]$$

وهذه العلاقة الأخيرة يُمكن كتابتها بالشّكل الآتي أيضاً:

$$a = \frac{\sum_{i=1}^{n} y_i}{n} - b\,\frac{\sum_{i=1}^{n} x_i}{n} = \overline{y} - b\,\overline{x}$$

أي إنَّ:

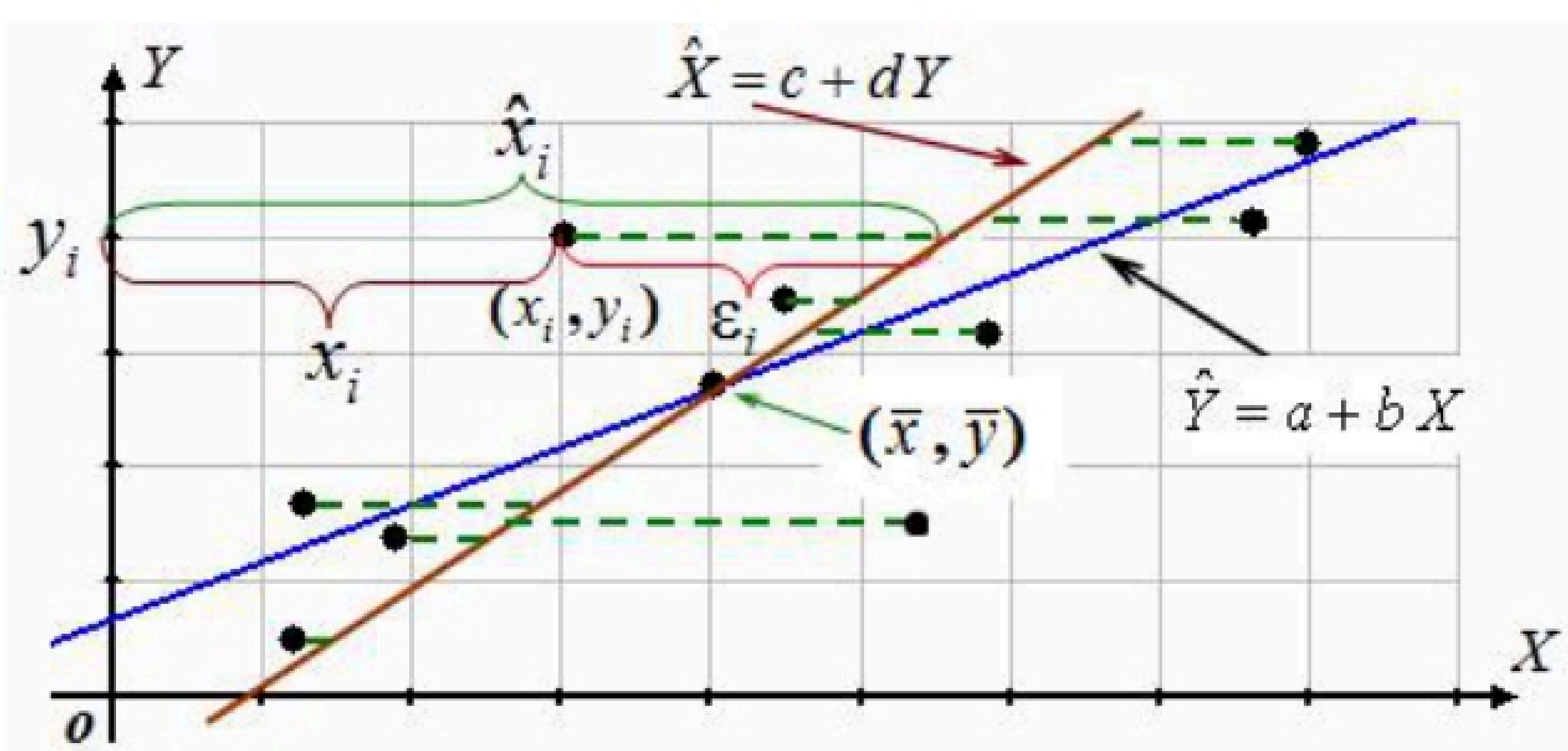
$$a = \overline{y} - b\,\overline{x} \qquad [3,15\text{-}b]$$

وهكذا تكون معادلة مستقيم الانحدار معيّنة تماماً عندما يكون الثابتان $a$ و $b$ قد تمّ حسابها من خلال العلاقتين [3–14-a] و [3–15-a] على الترتيب.

إذاً، وفقاً لهذه الطريقة يُنظر إلى هذا المستقيم (ويُدعى **مستقيم انحدار $Y$ على $X$**) على أنَّه أفضل مستقيم يمرّ بالنقاط (**أو بالقرب من النقاط**) الممثِّلة لهذه البيانات، وهذه العلاقة تساعدنا في تقدير (**أو التخمين عن**) القيمة $\hat{y}_o$ الموافقة لقيمة $x_o$ تقع **ضمن نطاق** بيانات المتغيِّر $X$ (**من الممكن $x_o$ ألّا تكون موجودة مع المشاهدات الأصل**)، وفي هذه الحالة تُدعى القيمة $\hat{y}_o$ بالقيمة المُقدَّرة لـ $Y$ والموافقة للقيمة المعلومة $x_o$.

### (٢ , ٢ , ٢ , ٣) معادلة مستقيم انحدار $X$ على $Y$. Equation of Straight Regression of $X$ on $Y$.

لقد لاحظنا أنَّ مُعامل بيرسون للارتباط الخطي بين المتغيِّرين $X$ و $Y$ هو نفسه بين $Y$ و $X$ لأنَّ $S_{X,Y} = S_{Y,X}$ (**مصفوفة التغاير للمتغيِّرين متناظرة دوماً**)، ولذلك قد يتبادر لذهن القارئ أن يكون مستقيم الانحدار $\hat{Y} = a + bX$ هو $\hat{X} = a + bY$ نفسه. لكن في الواقع هذين المستقيمين مختلفين عن بعضهما في الحالة العامّة، وذلك لأنَّ العلاقة [3–13-a] لا تتمتَّع بخاصية التناظر بـ $X$ و $Y$، ومن جهةٍ أخرى فلو عُدنا بذاكرتنا قليلاً للأخطاء $\varepsilon_i$ فقد حُسبت بناءً على الإسقاط الرأسي (**بشكل موازٍ لـ oY**) لـ $y$ على مستقيم الانحدار كما في الشكل (٣ , ١٦) السّابق، فحصلنا نتيجةً لذلك على المعادلة $\hat{Y} = a + bX$، وأمّا لو قمنا بحساب الأخطاء $\varepsilon_i$ بناءً على الإسقاط الأفقي (**الموازي لـ oX**) لـ $x$ على مستقيم الانحدار كما في الشكل الآتي:




الشكل (٣ , ١٧)

فعندئذ سنحصل على معادلة أخرى لمستقيم المربعات الصغرى من الشكل:

$$\hat{X} = c + dY \qquad [3,16\text{-}a]$$

ومن أجل $y_i$ قيمة مُحدَّدة لـ $Y$ يصبح لمعادلة هذا المستقيم العرض الآتي:

$$\hat{x}_i = c + d\,y_i \quad ; i \in \mathbb{N}_n \qquad [3,16\text{-}b]$$

علماً أنَّه من أجل كل $i \ni \mathbb{N}_n$ لدينا $\hat{x}_i$ هي القيمة المُقدَّرة من $y_i$ باستخدام المعادلة [3–16-a]، ومن ثمَّ يكون مقدار انزياح

$\hat{x}_i$ عن $x_i$ هو $\varepsilon_i := x_i - \hat{x}_i$ . علماً أنَّ $d$ تُحسب (وبشكل مماثل لما سبق شرحه) بالعلاقة الآتية:

$$d = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2} \qquad [3,17\text{-a}]$$

التي يمكن كتابتها بدلالة قيم مُعامل الارتباط الخطي $\mathbf{r}$ والانحراف المعياري لكل من $X$ و$Y$ على النحو الآتي أيضاً:

$$d = \frac{S_{X,Y}}{S_Y^2} = \frac{S_X}{S_Y}\,\mathbf{r} \qquad [3,17\text{-b}]$$

وهذه القيمة التي تُمثّل ميل مستقيم الانحدار تُدعى مُعامل انحدار $X$ على $Y$ ، وأمّا قيمة $c$ التي تُمثّل قيمة تقاطع هذا المستقيم مع المحور $\mathbf{oX}$ فإنّها تُحسب بالعلاقة الآتية:

$$c = \frac{\sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} y_i^2 - \left(\sum_{i=1}^{n} y_i\right)^2} \qquad [3,18\text{-a}]$$

ويمكن كتابتها بالشكل الآتي أيضاً:

$$c = \frac{\sum_{i=1}^{n} x_i}{n} - d\,\frac{\sum_{i=1}^{n} y_i}{n} = \overline{y} - d\,\overline{x}$$

أي إنَّ:

$$c = \overline{y} - d\,\overline{x} \qquad [3,18\text{-b}]$$

إنَّ مستقيم المربّعات الصغرى والمُعطى بالعلاقة [3–16-a] يُدعى مستقيم انحدار $X$ على $Y$ ، ومعادلته تساعدنا في تقدير (أو تخمين) القيمة $\hat{x}_o$ الموافقة لقيمة $y_o$ (من الممكن ألّا تكون موجودة مع المشاهدات) تقع ضمن نطاق بيانات المتغيِّر $Y$، وفي هذه الحالة تُدعى القيمة $\hat{x}_o$ بالقيمة المُقدَّرة لـ $X$ والموافقة للقيمة المعلومة $y_o$ .

### (٣,٢,٢,٣) ملاحظات:

١- إنَّ مستقيم الانحدار المستنتج وفقاً لطريقة المربّعات الصغرى يمرُّ بالنقطة التي إحداثييها $(\overline{x}\,,\,\overline{y})$، وذلك لأنَّه من العلاقة [3–15-b] لدينا $\overline{y} = a + b\,\overline{x}$ ، ومن ثَمّ ينتج أنَّ $\hat{y} - \overline{y} = b\,(x - \overline{x})$ ، ومن ثمَّ باستخدام العلاقة [3–14-b] ينتج لدينا العلاقة الآتية:

$$\hat{y} - \overline{y} = \frac{S_Y}{S_X}\mathbf{r}\left(x - \overline{x}\right) \;\Rightarrow\; \hat{y} = \overline{y} + \frac{S_Y}{S_X}\mathbf{r}\left(x - \overline{x}\right) \qquad [3,19]$$

وهذه العلاقة الأخيرة هي علاقة لمستقيم يمرّ بالنقطة $\left(\overline{x}\,,\,\overline{y}\right)$، أي إن النقطة $\left(\overline{x}\,,\,\overline{y}\right)$ تقع على مستقيم انحدار $Y$ على $X$ .

الآن، وبطريقة مماثلة يمكن إثبات أنَّ النقطة $\left(\overline{x}\,,\,\overline{y}\right)$ تقع على مستقيم انحدار $X$ على $Y$ أيضاً، ومن ثَمَّ يتقاطع مستقيم انحدار $Y$ على $X$ مع مستقيم انحدار $X$ على $Y$ في النقطة $\left(\overline{x}\,,\,\overline{y}\right)$، ونتيجةً لذلك ستنشأ بينهما زاوية قيمتها $0 \le \alpha \le \frac{\pi}{2}$ (انظر الشكل السّابق (٣,١٧))، وهذه الزاوية $\alpha$ ذات صلة وثيقة بالارتباط بين المتغيِّرين $X$ و$Y$، فكلّما كان الارتباط قوياً؛ صغرت قيمة زاوية التقاطع بحيث تصبح صفراً (أي إنَّ مستقيم انحدار $Y$ على $X$ سيكون منطبقاً على مستقيم انحدار $X$ على $Y$) عندما يصبح $\mathbf{r} = \pm 1$ وبشرط $b \neq 0$ ، وتكبر هذه الزاوية كلّما ضعف الارتباط بحيث تصبح قيمتها $\frac{\pi}{2}$ (مستقيما الانحدار متعامدان) عندما يصبح $\mathbf{r} = 0$ وبشرط $0 < S_X$ و

$S_Y > 0$، ولذلك فإنَّ اقتراب أطراف مستقيمي الانحدار $X$ على $Y$ وكذلك $Y$ على $X$ من بعضهما يدلُّل على قوة أكبر للارتباط الخطي بين المتغيِّرين (أو الظاهرتين) اللذين هما قيد الدراسة.

٢- من العلاقة [3,14-b] أو [3,17-b] يتضّح لنا أنّه يمكن تعيين معادلة مستقيم الانحدار إذا عُلمت قيمة مُعامل الارتباط $\mathbf{r}$ والانحراف المعياري مع المتوسطات لكل من بيانات المتغيِّرين $X$ و $Y$.

٣- من العلاقتين [3,14-b] أو [3,17-b] نلاحظ أنَّ حاصل ضرب مُعامل انحدار $Y$ على $X$ مع مُعامل انحدار $X$ على $Y$ يساوي قيمة مُعامل التحديد للمشاهدات، وذلك لأنَّ $\mathbf{r}\frac{S_Y}{S_X}\cdot\mathbf{r}\frac{S_X}{S_Y}=\mathbf{r}^2$، وبما $0 \le \mathbf{r}^2 \le 1$ فإنَّ ذلك يعني أنّه يمكن استخدام قيمة مُعامل التحديد للحكم على قوة الارتباط بين البيانات، فكلّما كانت قيمة هذا المُعامل أقرب إلى الواحد؛ دلَّ ذلك على قوة ارتباط أكبر بين البيانات، والعكس بالعكس صحيح.

٤- عندما يُطلب من برنامج إحصائي (مثل Minitab أو SPSS) تقديم مستقيم الانحدار لبيانات مُعطاة فإنَّه يَعرض لنا مستقيم انحدار $Y$ على $X$ فقط.

### (٣,٢,٢,٤) أمثلة

١- لتكن لدينا البيانات الآتية التي تمثّل درجات التحصيل النهائية لعشرة طلاب في مقرّرين دراسيين في الرياضيات:

الشكل (٣,١٩.أ)

| $i$ | المقرّر $X$ | المقرّر $Y$ |
|---|---|---|
| 1 | 75 | 69 |
| 2 | 82 | 85 |
| 3 | 65 | 55 |
| 4 | 90 | 90 |
| 5 | 77 | 80 |
| 6 | 60 | 50 |
| 7 | 55 | 50 |
| 8 | 87 | 90 |
| 9 | 77 | 72 |
| 10 | 45 | 60 |

ولنقم بدراسة الارتباط والانحدار الخطي لهذه البيانات.

من أجل ذلك لدينا:

$$\sum_{i=1}^{n} x_i = 713 \;\;\&\;\; \sum_{i=1}^{n} x_i^2 = 52751 \;\;\&\;\; \sum_{i=1}^{n} x_i\, y_i = 51804$$

$$\sum_{i=1}^{n} y_i = 701 \;\;\&\;\; \sum_{i=1}^{n} y_i^2 = 51395 \;\;\&\;\; (\overline{x},\overline{y}) = (71.3, 70.1)$$

ويفضَّل لتسهيل العمليات الحسابية للمقادير السابقة أن نبني الجدول الآتي:

الشكل (١٩ ,٣.ب)

| $i$ | $x_i$ | $x_i^2$ | $y_i$ | $y_i^2$ | $x_i \cdot y_i$ |
|---|---|---|---|---|---|
| 1 | 75 | 5625 | 69 | 4761 | 5175 |
| 2 | 82 | 6724 | 85 | 7225 | 6970 |
| 3 | 65 | 4225 | 55 | 3025 | 3575 |
| 4 | 90 | 8100 | 90 | 8100 | 8100 |
| 5 | 77 | 5929 | 80 | 6400 | 6160 |
| 6 | 60 | 3600 | 50 | 2500 | 3000 |
| 7 | 55 | 3025 | 50 | 2500 | 2750 |
| 8 | 87 | 7569 | 90 | 8100 | 7830 |
| 9 | 77 | 5929 | 72 | 5184 | 5544 |
| 10 | 45 | 2025 | 60 | 3600 | 2700 |
| sum | 713 | 52751 | 701 | 51395 | 51804 |

ومنه تكون قيمة مُعامل الارتباط الخطي للبيانات المُعطاة هي:

$$\mathbf{r}_P = \frac{10\times 51804 - 713\times 701}{\sqrt{10\times 52751 - 713^2}\,\sqrt{10\times 51395 - 701^2}} \approx 0.8773$$

وهذا يعني وجود ارتباط خطي إيجابي قوي جداً بين بيانات المتغيِّرين $X$ و $Y$.

أمّا لتعيين معادلة مستقيم انحدار $Y$ على $X$ فإنّنا نجد ما يلي:

$$b = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{10\times 51804 - 713\times 701}{10\times 52751 - 713^2} = 0.952$$

$$a = \frac{\sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2} = \frac{701\times 52751 - 713\times 51804}{10\times 52751 - 713^2} = 2.205$$

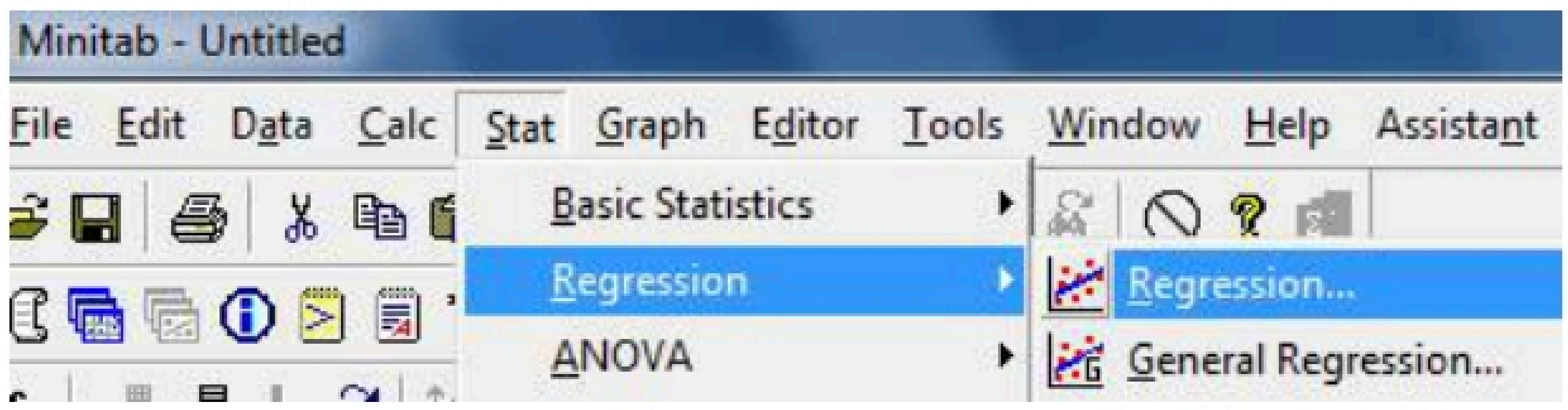
ومن ثَّم يكون لمعادلة مستقيم انحدار $Y$ على $X$ العرض الآتي:

$$\hat{Y} = 2.205 + 0.952\,X$$

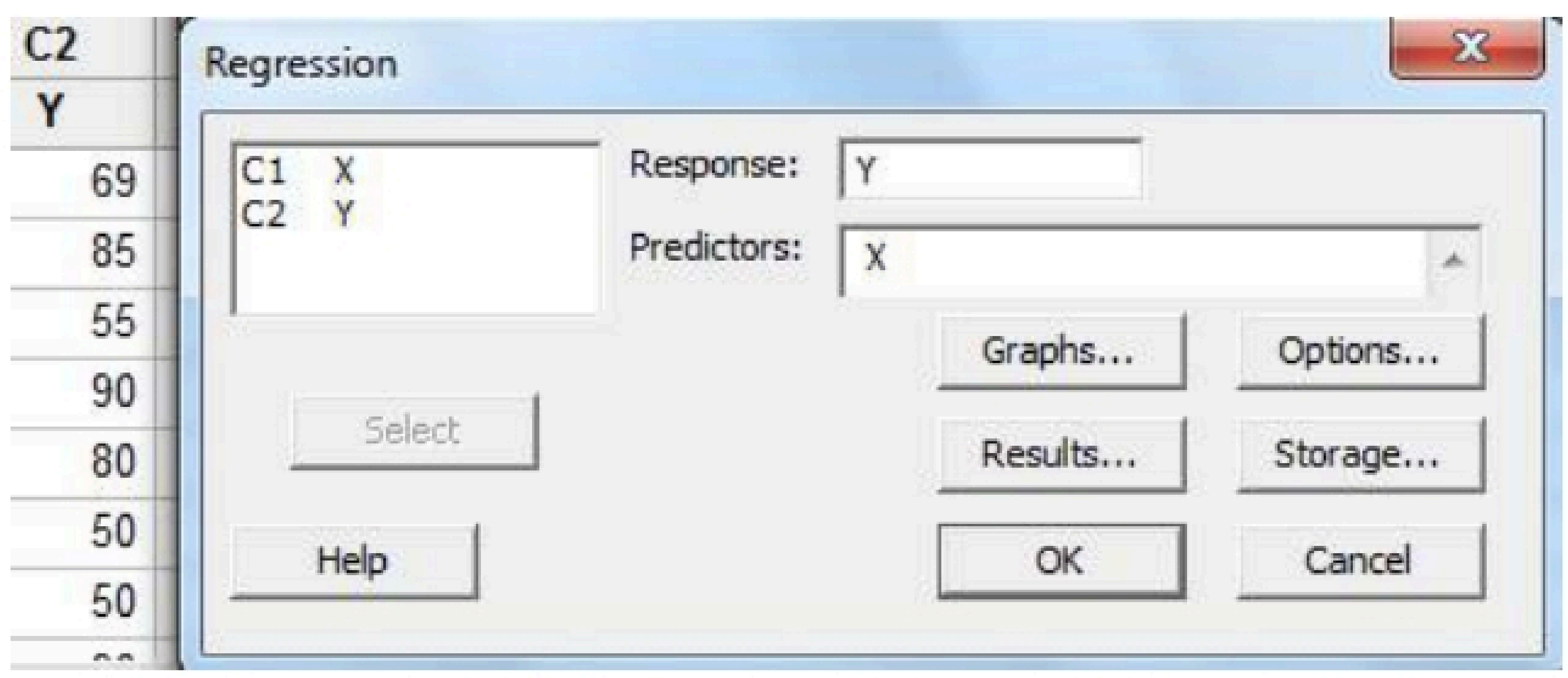
فلو أنّنا افترضنا أنَّ طالباً قد حصل على الدرجة 80 في المقرَّر $X$ ، فعندئذ ستكون الدرجة المقدّرة له في المقرّر $Y$ هي:

$$\hat{Y} = 2.205 + 0.952\,(80) = 78.365 \approx 78$$

من الممكن استخدام برنامج Minitab للحصول على هذه المعادلة السّابقة مع تمثيلها البياني على لوحة الانتشار وذلك باتباع الخطوات الموضَّحة بالصور الآتية:




ومن ثمَّ




فنحصل على تقرير منه المُعطيات الآتية:

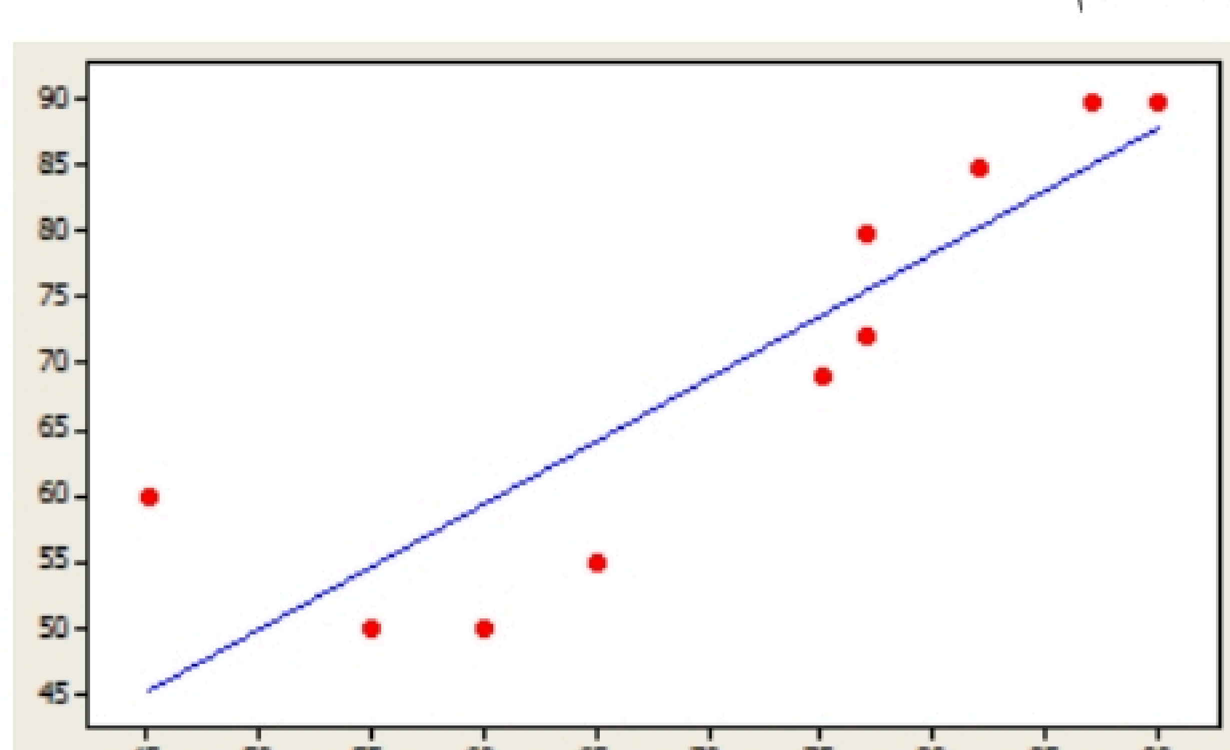
Regression Analysis: **Y** versus **X**

The regression equation is  Y = 2.2 + 0.952 X

وأمّا الرسم البياني لمستقيم انحدار $Y$ على $X$ فيتمّ وفقاً للخطوات الآتية:

Graph ⟹ Scatterplot ⟹ With Regression

فنحصل على العرض الآتي لهذا المستقيم:




الشكل (١٨, ٣.أ)

وأخيراً من أجل تعيين معادلة مستقيم انحدار $X$ على $Y$ لدينا:

$$d=\frac{n\sum_{i=1}^{n}x_iy_i-\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i}{n\sum_{i=1}^{n}y_i^2-\left(\sum_{i=1}^{n}y_i\right)^2}=\frac{10\times 51804-713\times 701}{10\times 51395-701^2}=0.808$$

$$c=\frac{\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i^2-\sum_{i=1}^{n}y_i\sum_{i=1}^{n}x_iy_i}{n\sum_{i=1}^{n}y_i^2-\left(\sum_{i=1}^{n}y_i\right)^2}=\frac{713\times 51395-701\times 51804}{10\times 51395-701^2}=14.636$$

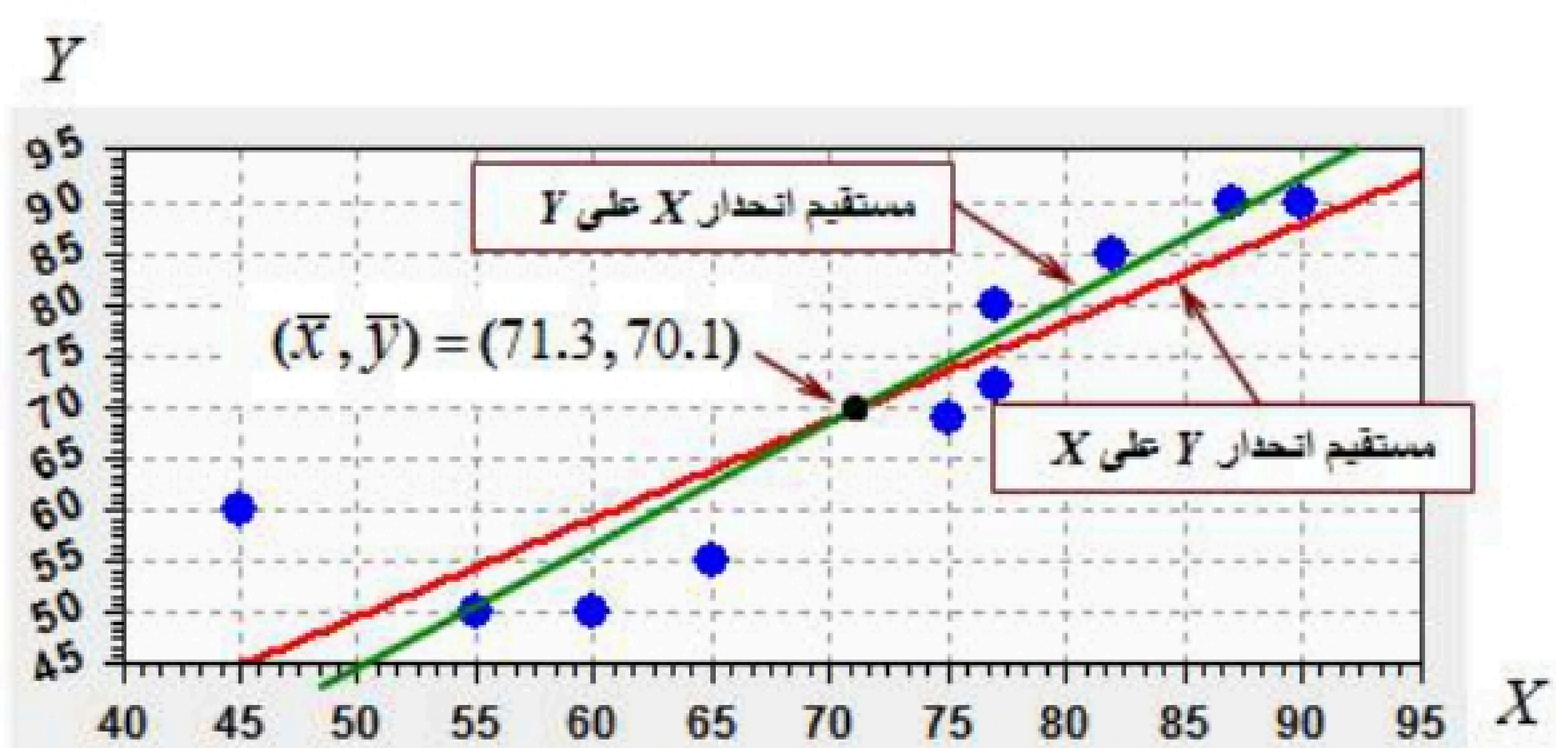
ومن ثمَّ يكون لمعادلة مستقيم انحدار $X$ على $Y$ العرض الآتي:

$$\hat{X}=14.636+0.808\,Y$$

فإذا افترضنا أنَّ طالباً قد حصل على الدرجة 66 في المقرَّر $Y$، فعندئذ ستكون الدرجة المقدّرة له في المقرّر $X$ هي:

$$\hat{X}=14.636+0.808\times 66=67.964\approx 68$$

وأمّا التمثيل البياني لمستقيمي الانحدار للمشاهدات المذكورة آنفًا سيكون لهما الشكل الآتي:




الشكل (١٨, ٣. ب).

لاحظ أنَّ الزاوية بين مستقيمي الانحدار صغيرة نسبياً وهذا يدلّل على القوة الكبيرة نسبياً للارتباط بين المتغيِّرين $X$ و $Y$.

٢- لتكن لدينا البيانات الآتية التي تمثّل المبيعات الشهرية لنوعين من السلع (الأطعمة $X$ والملابس $Y$) وعلى مدى عام كامل في سوق تجاري، علماً أنَّ القيم المقدَّمة مقدّرةً بـ 10,000 وحدة نقدية.

الشكل (٢٠, ٣.أ)

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $X$ | 13 | 33 | 100 | 85 | 16 | 45 | 75 | 31 | 42 | 77 | 36 | 85 |
| $Y$ | 12 | 45 | 20 | 66 | 33 | 45 | 22 | 60 | 33 | 66 | 55 | 45 |

الآن، ولدراسة الارتباط والانحدار الخطي لهذه المشاهدات سنقوم أولاً بحساب المجاميع الآتية:

$$\sum_{i=1}^{n}x_i \;\&\; \sum_{i=1}^{n}x_i^2 \;\&\; \sum_{i=1}^{n}y_i \;\&\; \sum_{i=1}^{n}y_i^2 \;\&\; \sum_{i=1}^{n}x_iy_i$$

والتي سنستخدمها لتعيين قيم مُعامل الارتباط الخطي ومعادلة مستقيمي الانحدار، ومن أجل ذلك سننشئ الجدول الآتي لتسهيل العمليات الحسابية:

الشكل (٢٠, ٣.ب)

| $i$ | $x_i$ | $x_i^2$ | $y_i$ | $y_i^2$ | $x_i \cdot y_i$ |
|---|---|---|---|---|---|
| 1 | 13 | 169 | 12 | 144 | 156 |
| 2 | 33 | 1089 | 45 | 2025 | 1485 |
| 3 | 100 | 10000 | 20 | 400 | 2000 |
| 4 | 85 | 7225 | 66 | 4356 | 5610 |
| 5 | 16 | 256 | 33 | 1089 | 528 |
| 6 | 45 | 2025 | 45 | 2025 | 2025 |
| 7 | 75 | 5625 | 22 | 484 | 1650 |
| 8 | 31 | 961 | 60 | 3600 | 1860 |
| 9 | 42 | 1764 | 33 | 1089 | 1386 |
| 10 | 77 | 5929 | 66 | 4356 | 5082 |
| 11 | 36 | 1296 | 55 | 3025 | 1980 |
| 12 | 85 | 7225 | 45 | 2025 | 3825 |
| sum | 638 | 43564 | 502 | 24618 | 27587 |

وهكذا يكون لدينا ما يلي:

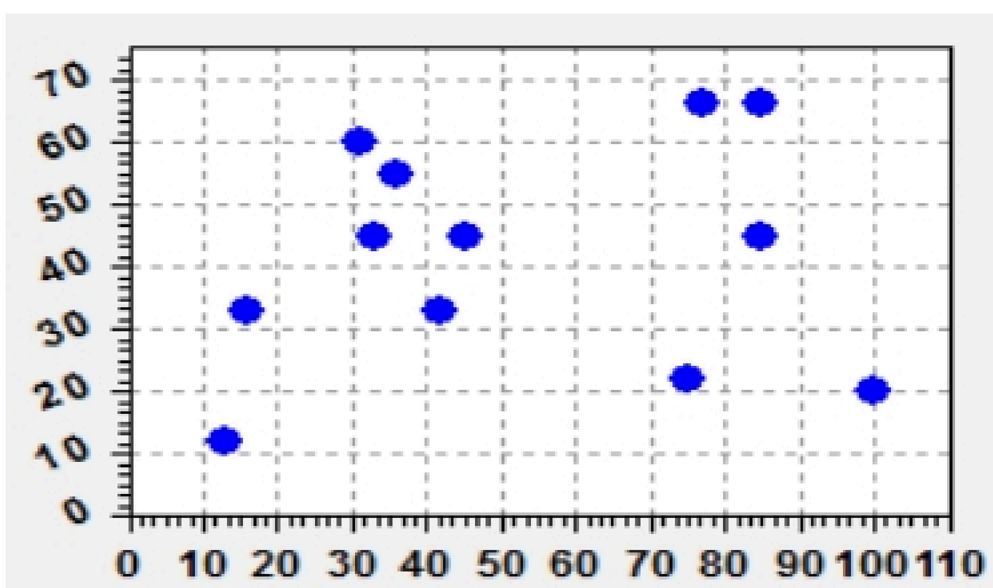
$$\sum_{i=1}^{n} x_i = 638 \;\; \& \;\; \sum_{i=1}^{n} x_i^2 = 43564 \;\; \& \;\; \sum_{i=1}^{n} y_i = 502 \;\; \& \;\; \sum_{i=1}^{n} y_i^2 = 24618 \;\; \& \;\; \sum_{i=1}^{n} x_i\, y_i = 27587$$

ومنه تكون قيمة مُعامل الارتباط الخطي لهذه البيانات تساوي:

$$\mathbf{r}_P = \frac{n \cdot \left(\sum_{i=1}^{n} x_i y_i\right) - \left(\sum_{i=1}^{n} x_i\right) \cdot \left(\sum_{i=1}^{n} y_i\right)}{\sqrt{n \cdot \left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2} \cdot \sqrt{n \cdot \left(\sum_{i=1}^{n} y_i^2\right) - \left(\sum_{i=1}^{n} y_i\right)^2}}$$

$$= \frac{12 \times 27587 - 638 \times 502}{\sqrt{12 \times 43564 - 638^2}\,\sqrt{12 \times 24618 - 502^2}} \approx 0.1519$$

وهذا يعني وجود ارتباط خطي إيجابي ضعيف جداً بين بيانات المتغيِّرين $X$ و$Y$، وهذا ما تأكده لوحة الانتشار الآتية للبيانات المقدَّمة:




الشكل (١٩, ٣.أ)

وأمّا لتعيين مستقيم انحدار $Y$ على $X$ فإنَّنا نجد:

$$b = \frac{12\times 27587 - 638\times 502}{12\times 43564 - 638^2} = 0.093 \qquad \& \qquad a = \frac{502\times 43564 - 638\times 27587}{12\times 43564 - 638^2} = 37.28$$

ومن ثمَّ يكون لمعادلة مستقيم انحدار $Y$ على $X$ الصيغة $\hat{Y} = 37.28 + 0.093X$، فلو أنّنا افترضنا أنَّ المتجر قد باع أطعمة بـ 800,000 وحدة نقديّة، فعندئذ سيكون المبلغ المقدّر لبيع الملابس هو:

$$\hat{Y} = 37.28 + 0.093\times 800000 = 74437.28$$

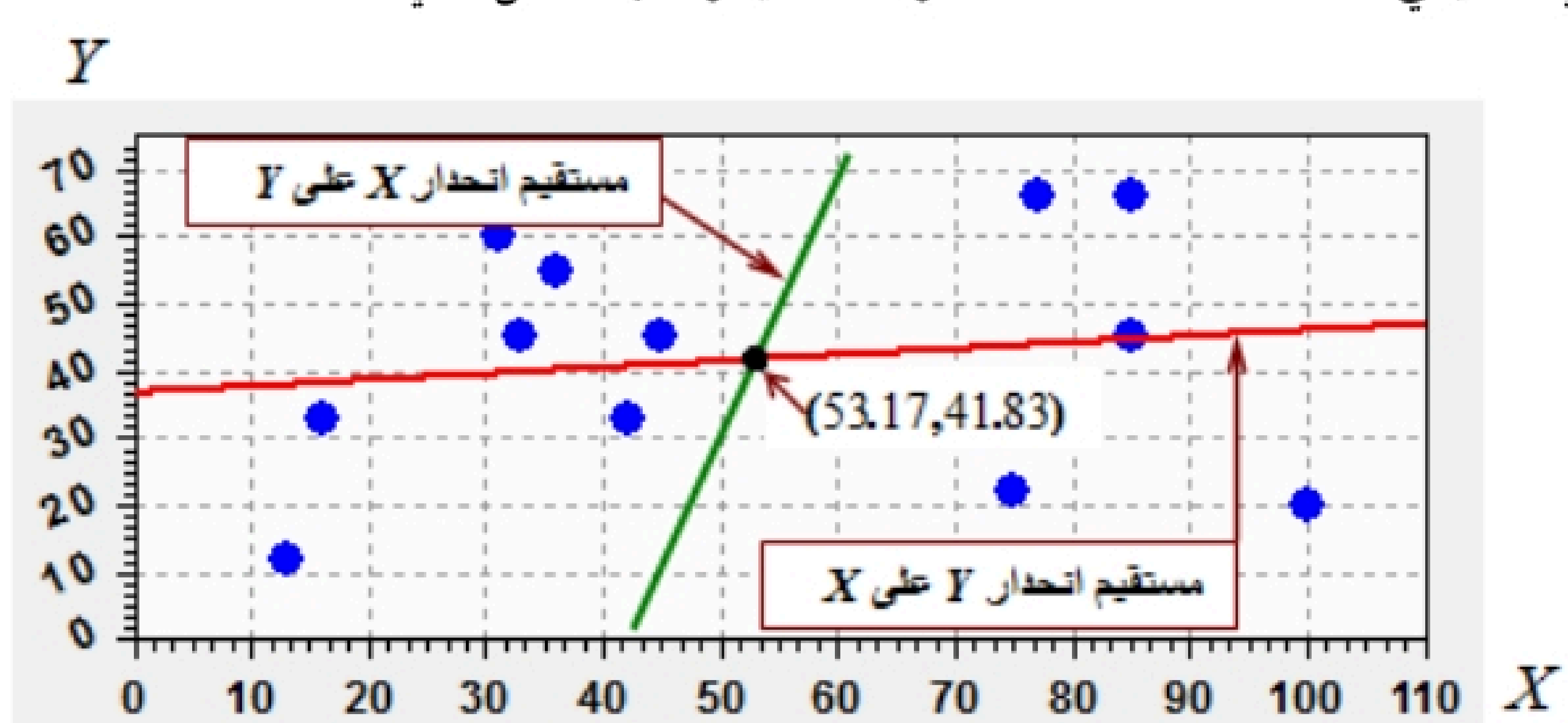
وأخيراً من أجل تعيين مستقيم انحدار $X$ على $Y$ لدينا:

$$d = \frac{10768}{43412} = 0.248 \qquad \& \qquad c = \frac{1857610}{43412} = 42.790$$

وبالتالي يكون لمعادلة مستقيم انحدار $X$ على $Y$ لهذه المشاهدات الصيغة $\hat{X} = 42.790 + 0.248Y$، فإذا افترضنا أنَّ المتجر قد باع ألبسةً بمبلغ 100,000 فعندئذ ستكون القيمة المقدّرة لمبيعات الأطعمة يساوي:

$$\hat{X} = 42.790 + 0.248\times 100000 = 24842.79$$

وأمّا التمثيل البياني لمستقيمي الانحدار للمشاهدات المذكورة آنفًا سيكون لهما الشكل الآتي:




الشكل (١٩, ٣. ب).

لاحظ كِبَر الزاوية بين مستقيمي الانحدار، وهي تشير إلى الضعف الشديد في الارتباط بين مبيعات الملابس والأطعمة.

### (٣, ٢, ٣) معادلتا مستقيما الانحدار لبيانات مجتمع إحصائي Equations of Straights Regression for Population Data

عادة لدى البحث في معادلة مستقيم الانحدار لبيانات مجتمع إحصائي ينطلق المرء من الفرض أنَّ حجم المجتمع غير منته أو كبير جداً، فإن كان حجم المجتمع الإحصائي منتهياً ويساوي $N$، فعندئذ بفرض أنَّ $(x_1, y_1)$، $(x_2, y_2)$، ...، و $(x_N, y_N)$ هي بيانات هذا المجتمع والنّاتجة عن إخضاع عناصره لمتغيِّرين $X$ و$Y$، فإنَّه يتمّ تعيين معادلة مستقيم الانحدار لبيانات هذا المجتمع كما في حالة العينات (وكأن المجتمع هنا هو عيِّنة كبيرة الحجم)، ولهذا السبب تقوم معظم البرامج الإحصائية التي تحتوي على فقرة تحليل الانحدار بالتعامل مع البيانات المقدّمة على أنّها بيانات عيِّنة، ومن ثمَّ تقدّم معادلة انحدار $Y$ على $X$ من أجل العينات فقط.

أمّا إذا كان حجم المجتمع غير منتهٍ فعندئذ لتعيين النموذج الرياضياتي لمعادلة مستقيم الانحدار سنحتاج إلى معارف من نظرية الاحتمالات (سنأتي على دراسة الاحتمالات في فصول قادمة)، ولكن سنقدّم عرضاً موجزاً لهذه المسألة حيث سنناقش تعيين مستقيم انحدار $Y$ على $X$ وكذلك مستقيم انحدار $X$ على $Y$.

### (١, ٣, ٢, ٣) معادلة مستقيم انحدار $Y$ على $X$:

إنَّ لمعادلة مستقيم انحدار $Y$ على $X$ لمجتمع إحصائي لها العرض الآتي:

$$\hat{Y} = \alpha + \beta X + \varepsilon \qquad [3,20\text{-a}]$$

علماً أنَّ $\alpha$ (تمثِّل قيمة تقاطع مستقيم الانحدار مع المحور oY) و $\beta$ (تمثِّل قيمة الميل لهذا المستقيم) هما ثابتان حقيقيان يُطلب تعيينهما، ويُدعيان مَعْلمي مستقيم الانحدار. أمّا $\varepsilon$ فإنَّه يُدعى مُركَّبة الخطأ العشوائي Random Error Component وله خصائص مُحدَّدة يجب أن يستوفيها (هذه الخصائص تتعلَّق باستقلال الأخطاء وتوزيعها الاحتمالي، ولهذا لن نخوض في تفاصيلها).

الآن، ومن أجل $x_i$ قيمة مُحدَّدة لـ $X$ فإنَّ قيمة $\hat{y}_i$ المُقدِّرة لـ $y_i$ تُحسب وفقاً للصيغة الآتية:

$$\hat{y}_i = \alpha + \beta x_i + \varepsilon_i \qquad [3,20\text{-b}]$$

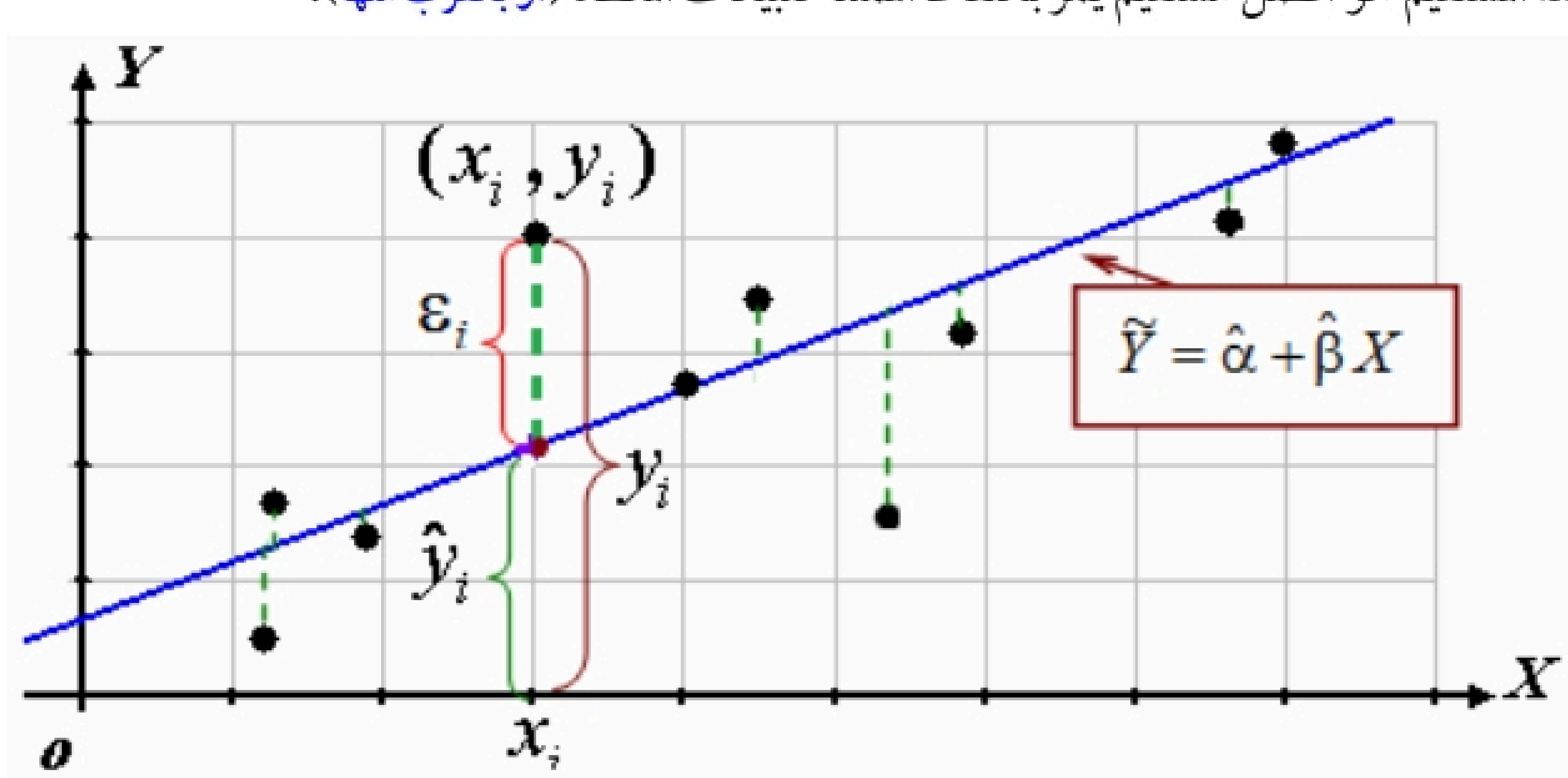
ومن ثَمَّ الأخطاء $\varepsilon_i = y_i - \hat{y}_i$ من أجل كل $i$ هي قيم عشوائية، وبما أنَّ $y_i$ قيم ثابتة فإنَّ القيم $\hat{y}_i$ ستصبح قيماً عشوائية أيضاً، وهكذا نلاحظ أنَّ معادلة مستقيم الانحدار للمجتمع يمكن أن تُجزأ إلى مجموع مُركبتين هما:

$$\hat{Y} = \underbrace{\alpha + \beta X}_{\text{Systematic Component}} + \underbrace{\varepsilon}_{\text{Random Error Component}}$$

١- المُركَّبة $\tilde{Y} := \hat{\alpha} + \hat{\beta} X$، وهي مركَّبة غير عشوائية تُدعى المُركَّبة النظامية (أو الرتيبة) Systematic Component حيث لدينا في هذه الحالة $\hat{\alpha}$ هو مُقدِّر المَعْلَمة $\alpha$، وأمّا $\hat{\beta}$ فهو مقدِّر المَعْلَمة $\beta$.

٢- المركَّبة $\varepsilon$ تُدعى المُركَّبة العشوائية كما ذكرنا ذلك سابقاً.

نلاحظ ممّا سبق أنَّه لتعيين معادلة هذا المستقيم فإنَّ المسألة قد آلت إلى تعيين قيم مقدِّرات الوسطاء $\alpha$ و $\beta$ لهذا المستقيم، وبحيث يكون هذا المستقيم هو أفضل مستقيم يمرّ بالنقاط الممثِّلة للبيانات المُعطاة (أو بالقرب منها).




الشكل (٣,٢٠)

الآن لتقدير قيم $\alpha$ و $\beta$ نقوم بسحب عيِّنة عشوائية (وليكن بحجم $n$) من المجتمع ونخضعها للمتغيِّرين $X$ و $Y$، فنحصل على البيانات $(x_1, y_1)$، $(x_2, y_2)$، ...، و $(x_n, y_n)$، ومن ثَمَّ نقوم بتعيين معادلة مستقيم انحدار $Y$ على $X$ وفقاً لطريقة المربّعات الصغرى فيكون لهذه المعادلة العرض الآتي:

$$\tilde{Y} = \hat{\alpha} + \hat{\beta} X \qquad [3,21]$$

وبعد ذلك تُضاف المركِّبة العشوائية $\varepsilon$ إلى النموذج فنحصل على العرض الآتي للنموذج المُقدَّرة لمعادلة مستقيم انحدار $Y$ على $X$:

$$\hat{Y} = \hat{\alpha} + \hat{\beta} X + \varepsilon \qquad [3,22]$$

الآن بملاحظة الصعوبة البالغة في تعيين القيم المقدَّرة $\hat{y}_i$ من خلال العلاقة [3,22] (أو من خلال العرض البياني الممثِّل لها) فإنَّ الدارس لا يهتمّ بالقيمة الدقيقة لـ $\hat{y}_i$ وإنّما يَلجأ إلى تعيين فترة تنتمي إليها هذه القيمة باحتمال معلوم، وهذا بدوره يقود إلى ما يُعرف باسم فترات الثقة لمستقيم الانحدار التي تقع خارج إطار هذا الفصل لاعتمادها على علم العشوائيات (نظرية الاحتمالات والإحصاء الرياضياتي).

### (٣, ٢, ٣, ٢) معادلة مستقيم انحدار $X$ على $Y$:

من خلال حوار مماثل تماماً لما سبق يكون لمعادلة مستقيم انحدار $X$ على $Y$ العرض الآتي:

$$\hat{X} = \gamma + \delta Y + \varepsilon \qquad [3,23\text{-a}]$$

ومن أجل $y_i$ قيمة مُحدَّدة لـ $Y$ يصبح لمعادلة هذا المستقيم العرض الآتي:

$$\hat{x}_i = \gamma + \delta y_i + \varepsilon_i \qquad [3,23\text{-b}]$$

علماً أنَّ $\gamma$ (تمثِّل قيمة تقاطع مستقيم الانحدار مع المحور **oX**) و $\delta$ (تمثِّل قيمة الميل لهذا المستقيم) ثوابت حقيقية يُطلب تعيينها، وتُدعى مَعَالِم مستقيم الانحدار، وأمّا $\varepsilon$ فهي **مُركِّبة الخطأ العشوائي**. ومن ثمَّ سيكون الخطأ $\varepsilon_i = x_i - \hat{x}_i$ المرافق لتقدير قيمة $x_i$ الموافقة لـ $y_i$ عند استخدام معادلة المستقيم [3,23-b] هي قيم عشوائية أيضاً.

الآن، وبعد تعيين $\hat{\gamma}$ و $\hat{\delta}$ كمقدِّرين للمعلمتين $\gamma$ و $\delta$ على الترتيب يصبح للمعادلة المُقدِّرة لمستقيم انحدار $X$ على $Y$ العرض الآتي:

$$\hat{Y} = \hat{\gamma} + \hat{\delta} X + \varepsilon \qquad [3,24]$$

وأمّا بقية الأفكار التي نوقشت في الحالة السابقة فإنَّها تُناقش هنا بشكل مماثل تماماً.

### (٣, ٢, ٣, ٣) ملاحظة

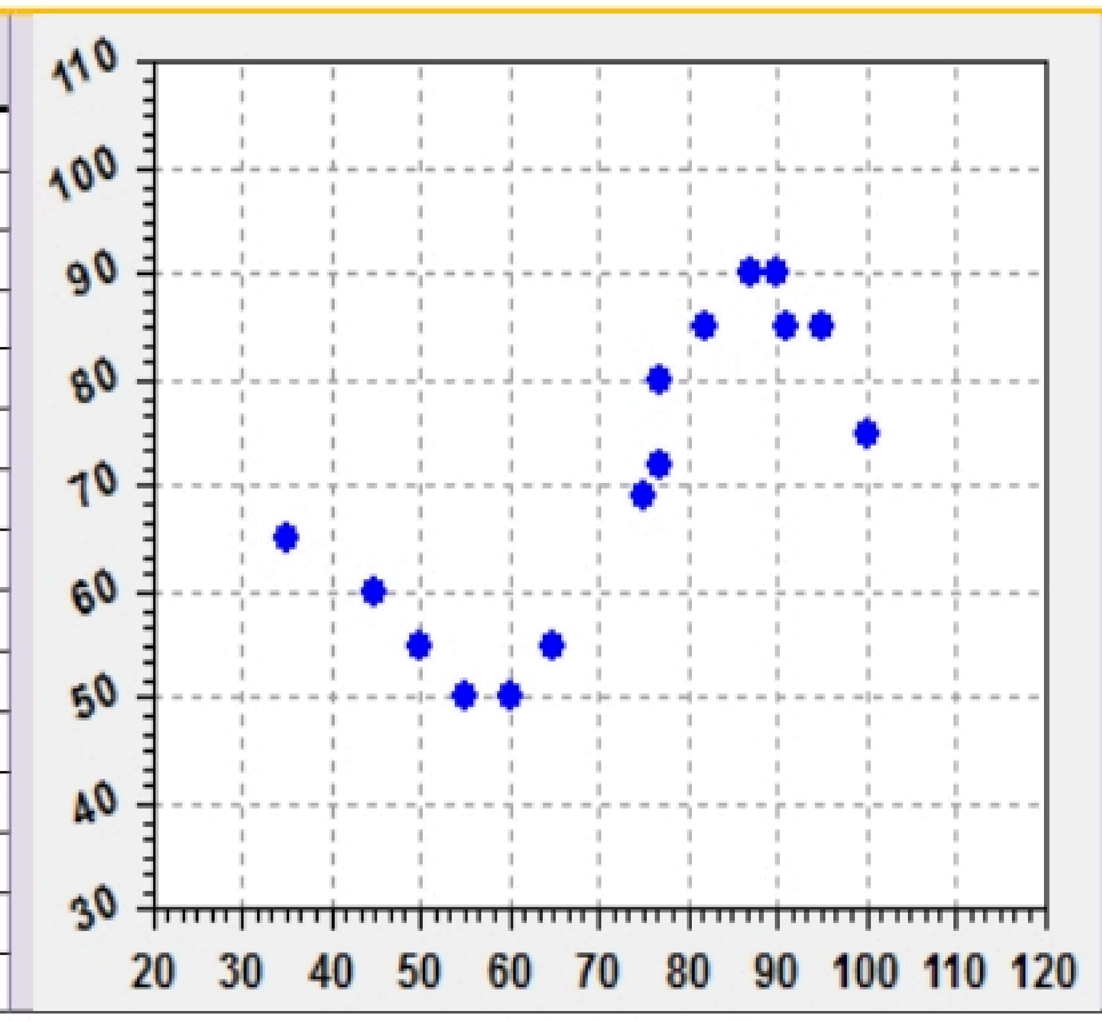
نودّ التأكيد على عدم المبالغة في استخدام معادلة مستقيم الانحدار في إيجاد قيمة مُقدَّرة $\hat{y}_o$ موافقة لقيمة $x_o$ كيفية، بمعنى أنَّه من غير المحبذ (إن لم يكن من غير الجائز في بعض الحالات) استخدام المعادلة [3,20-a] (أو المعادلة [3,23-a]) في التنبؤ بقيمة $\hat{y}_o$ المقابلة لقيمة $x_o$ الواقعة خارج نطاق بيانات المتغيِّر $X$ (وفي هذه الحالة تُدعى القيمة $\hat{y}_o$ بالقيمة المتنبأ بها لـ $Y$ والموافقة لـ $x_o$)، وذلك لأنَّه من أجل القيم التي تكون خارج الفترة التي تقع فيها مشاهدات $X$ قد تكون العلاقة بين قيم $X$ و$Y$ غير خطّيّة، ومن ثمَّ تصبح النتائج التي نحصل عليها كقيم مُقدَّرة ليست صحيحة أو غير دقيقة على الأقل، والمثال الآتي يوضِّح لنا ذلك.

### (٣, ٢, ٣, ٤) مثال

لتكن لدينا المشاهدات $(x_i , y_i)$ مع $i \in \mathbb{N}_{15}$ المقدّمة في الجدول (١٨ ,٣ .أ) الآتي، ولنقم بتمثيل هذه البيانات على لوحة الانتشار ومن ثمَّ تعيين معادلة مستقيم انحدار $Y$ على $X$ لهذه المشاهدات، وبعد ذلك مقارنة نتائج التنبؤ بهذا المستقيم مع المنحنى الموافق لمسار هذه البيانات.

الجدول (٢١,٣.أ)

| $i$ | $x_i$ | $y_i$ |
|---|---|---|
| 1 | 75 | 69 |
| 2 | 82 | 85 |
| 3 | 65 | 55 |
| 4 | 90 | 90 |
| 5 | 77 | 80 |
| 6 | 60 | 50 |
| 7 | 55 | 50 |
| 8 | 87 | 90 |
| 9 | 91 | 85 |
| 10 | 77 | 72 |
| 11 | 50 | 55 |
| 12 | 45 | 60 |
| 13 | 95 | 85 |
| 14 | 100 | 75 |
| 15 | 35 | 65 |




الشكل (٢١,٣.أ)

فنجد للبيانات المُعطاة العرض الانتشاري المقدَّمة في الشكل السّابق (٢١,٣.أ) (حيث يُلاحظ لمسار البيانات شكل علاقة غير خطيّة)، وأمّا من أجل حساب الأمثال $a$ و$b$ في معادلة مستقيم الانحدار فإنّنا سنقدِّم الجدول الآتي لتسهيل الحسابات:

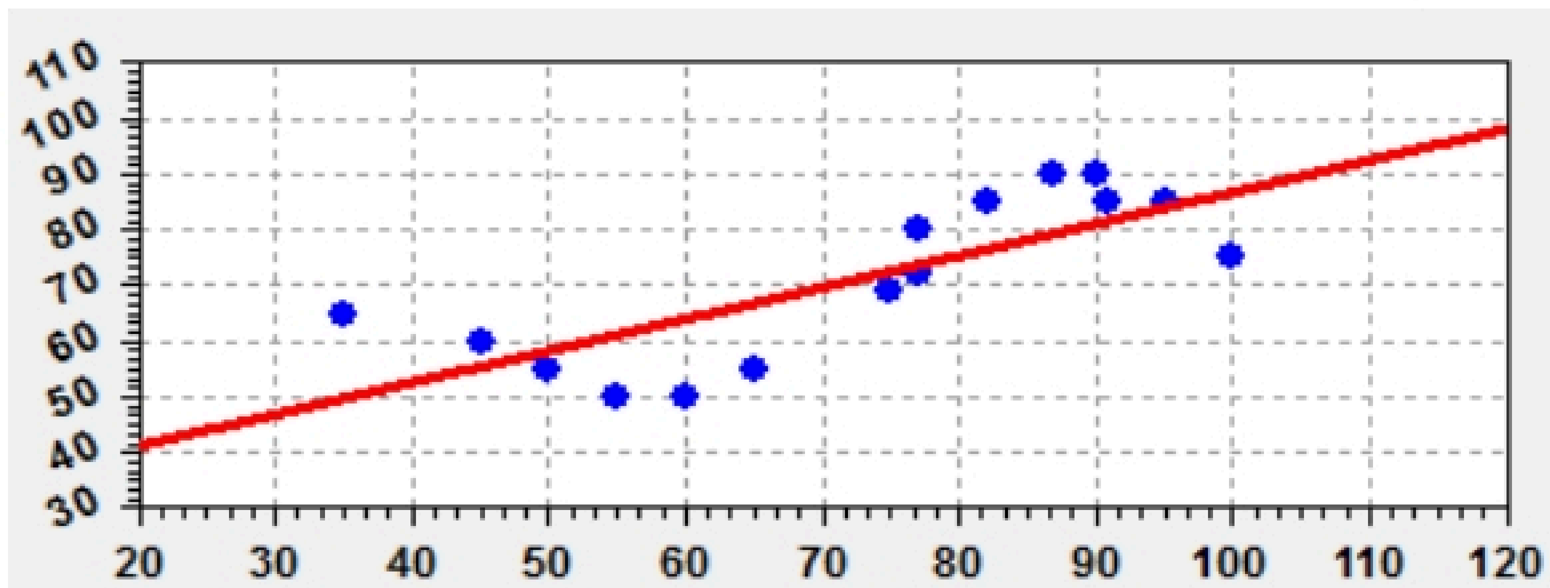
الجدول (٢١,٣.ب).

| $i$ | $x_i$ | $x_i^2$ | $y_i$ | $y_i^2$ | $x_i \cdot y_i$ |
|---|---|---|---|---|---|
| 1 | 75 | 5625 | 69 | 4761 | 5175 |
| 2 | 82 | 6724 | 85 | 7225 | 6970 |
| 3 | 65 | 4225 | 55 | 3025 | 3575 |
| 4 | 90 | 8100 | 90 | 8100 | 8100 |
| 5 | 77 | 5929 | 80 | 6400 | 6160 |
| 6 | 60 | 3600 | 50 | 2500 | 3000 |
| 7 | 55 | 3025 | 50 | 2500 | 2750 |
| 8 | 87 | 7569 | 90 | 8100 | 7830 |
| 9 | 91 | 8281 | 85 | 7225 | 7735 |
| 10 | 77 | 5929 | 72 | 5184 | 5544 |
| 11 | 50 | 2500 | 55 | 3025 | 2750 |
| 12 | 45 | 2025 | 60 | 3600 | 2700 |
| 13 | 95 | 9025 | 85 | 7225 | 8075 |
| 14 | 100 | 10000 | 75 | 5625 | 7500 |
| 15 | 35 | 1225 | 65 | 4225 | 2275 |
| **sum** | **1084** | **83782** | **1066** | **78720** | **80139** |

فنجد باستخدام العلاقتين [3–14-a] و [3–15-a] أنَّ:

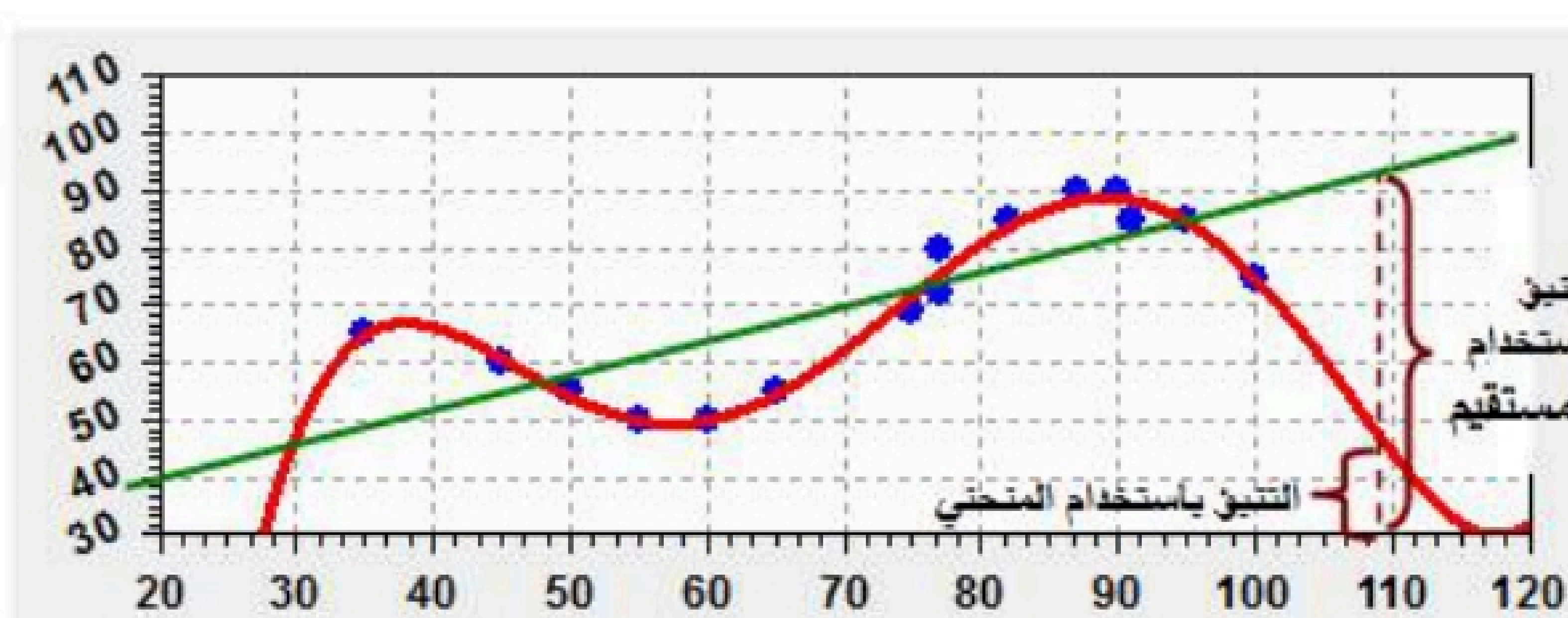
$$b = \frac{15\times 80139 - 1084\times 1066}{15\times 83782 - 1084^2} \approx 0.57 \qquad \& \qquad a = \frac{1066\times 83782 - 1084\times 80139}{15\times 83782 - 1084^2} \approx 29.886$$

ومن ثمَّ يكون لمعادلة مستقيم انحدار $Y$ على $X$ الصيغة $\hat{Y} = 29.886 + 0.57X$، ورسمها البياني كما في الشكل الآتي:



الشكل (٢١,٣.ب)

الآن لو قمنا بالبحث عن منحنى مناسب لهذه البيانات فإنَّنا نجد باستخدام برنامج Curve Expert أنَّ كثيرة حدود من الدرجة الخامسة تُعطينا توفيقاً جيداً للبيانات (انظر الشكل الآتي).




شكل (٢١,٣ج)

فلو أردنا التنبؤ بالقيمة المقابلة لـ $x_o = 110$ فإنَّنا سنلاحظ الخطأ الكبير نسبياً بين القيمة المتنبأ بها باستخدام مستقيم الانحدار والقيمة المتنبأ بها باستخدام منحنى كثيرة الحدود من الدرجة الخامسة الراسم للمسار الحقيقي (تقريباً) للبيانات، وهكذا نخلص إلى أنَّه إذا كانت البيانات تتوزَّع على شكل علاقة غير خطيِّة، فإنَّنا بالتأكيد سنحصل على الخط المستقيم الموَّفق للمشاهدات، ولكن هذا الخط قد لا يُعطينا نتائج ذات مغزى في عمليات التنبؤ.

### (٣,٢,٤) دقَّة التقديرات Accuracy of Estimates

لقد لاحظنا أنَّ القيم الفعليَّة لن تقع جميعها على خط الانحدار، بل سيقع بعضها على مسافات متفاوتة عنه، ومن ثمَّ سيوجد اختلاف بين القيم المُقدَّرة بمعادلة مستقيم الانحدار والقيم الفعليَّة (المشاهدة)، ولهذا نحتاج إلى مقياس يُبيِّن لنا بدقَّة مدى ملاءمة هذا المستقيم للنقاط الممثِّلة للبيانات.

إنَّ أحد هذه المقاييس يعتمد على تشتت القيم المشاهدة عن مستقيم الانحدار الذي يُعطينا ما يُسمَّى بـ **الخطأ المعياري للتقدير** Standard Error of the Estimate، والذي يُستخدم كمقياس لدرجة الدقَّة في تقدير القيم المُقدَّرة (أو المتنبأ بها في مسائل التنبؤ)، وبما أنَّ الانحراف المعياري هو أفضل مقاييس التشتت لمجموعة بيانات (كما ذكرنا ذلك في فصل سابق) فإنَّه يمكن استخدام مقياس مكافئ له من أجل انحرافات القيم الحقيقية عن مستقيم الانحدار.

نعلم أنَّ الانحراف المعياري هو الجذر التربيعي الموجب لمتوسط مربّعات انحرافات البيانات عن متوسطها، ومن ثمَّ يمكننا حساب الانحراف المعياري حول مستقيم الانحدار باستخدام الأسلوب نفسه، ولكن مع الأخذ بالحسبان أنَّه لدينا بيانات من نوع

مختلف (**ثنائيات**) وأنَّ مستقيم الانحدار هو الذي سيقوم بدور الممثِّل للقيم الوسطى لهذه البيانات حيث لاحظنا أنَّ ذلك يؤدي إلى ظهور صيغ قد تكون مختلفة بعض الشيء عن تلك التي للبيانات المفردة.

### (١, ٤, ٢, ٣) الخطأ المعياري لتقدير مستقيم الانحدار

لتكن $(x_1,y_1)$، $(x_2,y_2)$، ...و $(x_n,y_n)$ بيانات **لعيِّنة** أخضعت لمتغيِّرين $X$ و $Y$. عندئذ:

١- بفرض أنَّ $\hat{Y} = a + bX$ هي معادلة مستقيم انحدار $Y$ على $X$، فإنَّ **الخطأ المعياري** لتقدير هذا المستقيم يُعرَّف من خلال العلاقة:

$$S_{est\hat{Y}} := \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n-2}} \qquad [3,25\text{-}a]$$

علماً أنَّ $\hat{y}_i$ هي القيمة المقدَّرة لـ $Y$ والموافقة لـ $x_i$ بوساطة مستقيم الانحدار، وأمّا $y_i$ فهي القيمة الفعليَّة (**المشاهدة**)، ولكن بملاحظة صعوبة استخدام العلاقة السابقة لحساب الخطأ المعياري لتقدير مستقيم الانحدار، فإنَّه يمكن استخدام قيم المشاهدات مباشرة لحساب هذا الخطأ وذلك من خلال العلاقة الآتية:

$$S_{est\hat{Y}} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \overline{y})^2 - b\sum_{i=1}^{n}(x_i - \overline{x})\cdot(y_i - \overline{y})}{n-2}} \qquad [3,25\text{-}b]$$

مع العلم أنَّ $\overline{x}$ هو المتوسط للبيانات $x_1$، $x_2$، ... و $x_n$، وكذلك $\overline{y}$ هو المتوسط للبيانات $y_1$، $y_2$، ... و $y_n$. أو باستخدام قيمة الانحراف المعياري $S_Y$ ومعامل الارتباط الخطي $\mathbf{r}$ وذلك بوساطة العلاقة المكافئة الآتية:

$$S_{est\hat{Y}} = S_Y\sqrt{\frac{n-1}{n-2}\left(1-\mathbf{r}^2\right)} \qquad [3,25\text{-}c]$$

والتي ينتج عنها أنَّ:

$$\mathbf{r} = \sqrt{1 - \frac{(n-2)S_{est\hat{Y}}^2}{(n-1)S_Y^2}} \qquad [3,26]$$

وهذه العلاقة تبيّن وجود علاقة جبرية بين الارتباط والتباين حول خط الانحدار، إذ إنَّ هذه الأخيرة توضِّح أنَّ مُعامل الارتباط سيكون كبيراً إذا كان التشتت حول مستقيم الانحدار صغيراً والعكس بالعكس صحيح أيضاً.

٢- بفرض أنَّ $\hat{X} = c + dY$ هي معادلة مستقيم انحدار $X$ على $Y$، فإنَّ **الخطأ المعياري** لتقدير هذا المستقيم يُعرَّف من خلال العلاقة الآتية:

$$S_{est\hat{X}} := \sqrt{\frac{\sum_{i=1}^{n}\left(x_i - \hat{x}_i\right)^2}{n-2}} \qquad [3,27\text{-}a]$$

أو باستخدام العلاقة المكافئة الآتية:

$$S_{est\hat{X}} = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \overline{x})^2 - d\sum_{i=1}^{n}(x_i - \overline{x})\cdot(y_i - \overline{y})}{n-2}} \qquad [3,27\text{-}b]$$

علماً أنَّ $\hat{x}_i$ هي القيمة المقدَّرة لـ $X$ والموافقة لـ $y_i$ بوساطة مستقيم انحدار $X$ على $Y$، وأمَّا $x_i$ فهي القيمة الفعليَّة المقابلة لـ $y_i$.

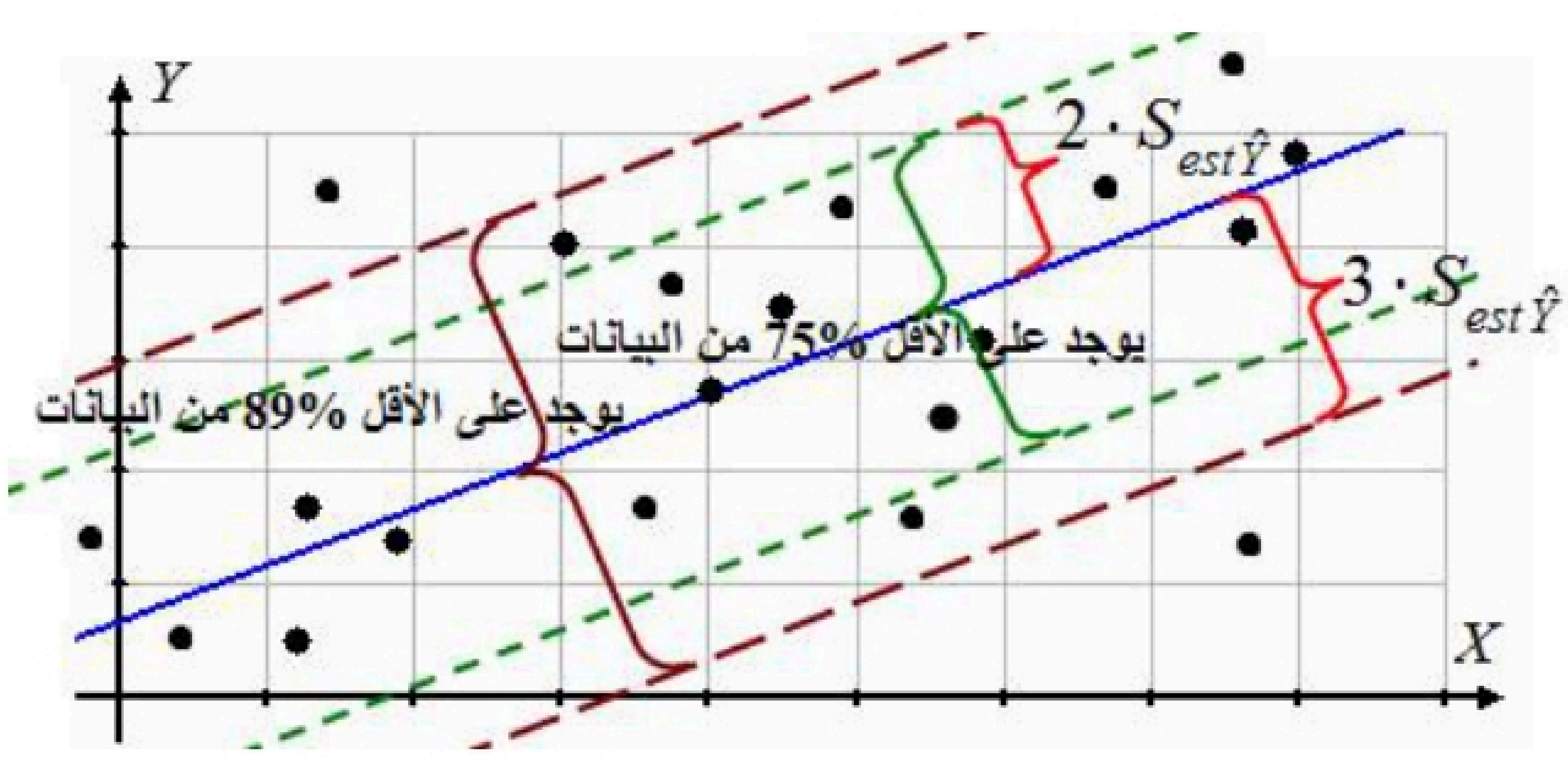
كذلك يمكن حساب الخطأ المعياري لتقدير هذا المستقيم باستخدام قيمة الانحراف المعياري $S_Y$ ومعامل الارتباط الخطي $\mathbf{r}$ وذلك بوساطة العلاقة المكافئة الآتية:

$$S_{est\,\hat{X}} = S_X \sqrt{\frac{n-1}{n-2}\left(1-\mathbf{r}^2\right)} \qquad [3,27\text{-}c]$$

والتي ينتج عنها أنَّ:

$$\mathbf{r} = \sqrt{1-\frac{(n-2)S_{est\,\hat{X}}}{(n-1)S_X}} \qquad [3,28]$$

وهكذا نلاحظ التشابه الكبير بين معادلة الخطأ المعيّاري لتقدير مستقيم الانحدار ومعادلة الانحراف المعياري لبيانات مُعطاة، وذلك لأنَّ كلاً منها يقيس التشتت حول قيمة متوسطة، فالأول يقيس انحرافات قيم المشاهدات عن القيم النظرية التي كان يجب أن يأخذها لو اتبعت هذه المشاهدات معادلة مستقيم الانحدار، وأمَّا الثاني فإنَّه يقيس انحرافات القيم حول نقطة معيَّنة هي المتوسط، ومن ثمَّ سيكون لكل منها الخصائص التي للانحراف المعياري نفسها، وبناءً على ما سبق يمكن تطبيق قاعدة تشبيشيف التجريبية من أجل مستقيم الانحدار أيضاً بحيث إنَّه لو قمنا برسم خطين مستقيمين موازيين لمستقيم الانحدار (**ولكن من جهتين مختلفتين ويبعدان عنه بمقدار انحرافين معياريين**) لوجب وجود 75٪ من النقاط الممثِّلة للمشاهدات على الأقل ضمن هذين المستقيمين. أمَّا إذا كانا يبعدان بمقدار ثلاثة انحرافات معيارية عن مستقيم الانحدار، فإنَّه يفترض وجود 89٪ على الأقل من النقاط الممثِّلة المشاهدات ضمن هذين المستقيمين، وهكذا .... (**انظر الشكل التوضيحي الآتي**).




الشكل (٣,٢٢)

### (٣,٢,٤,٢) ملاحظات:

١- إنَّ الخطأ المعياري لتقدير مستقيمي الانحدار $S_{est\hat{Y}}$ و $S_{est\,\hat{X}}$ يُدعيان **متوسط مربَّع الخطأ** Mean Squared Error أيضاً.

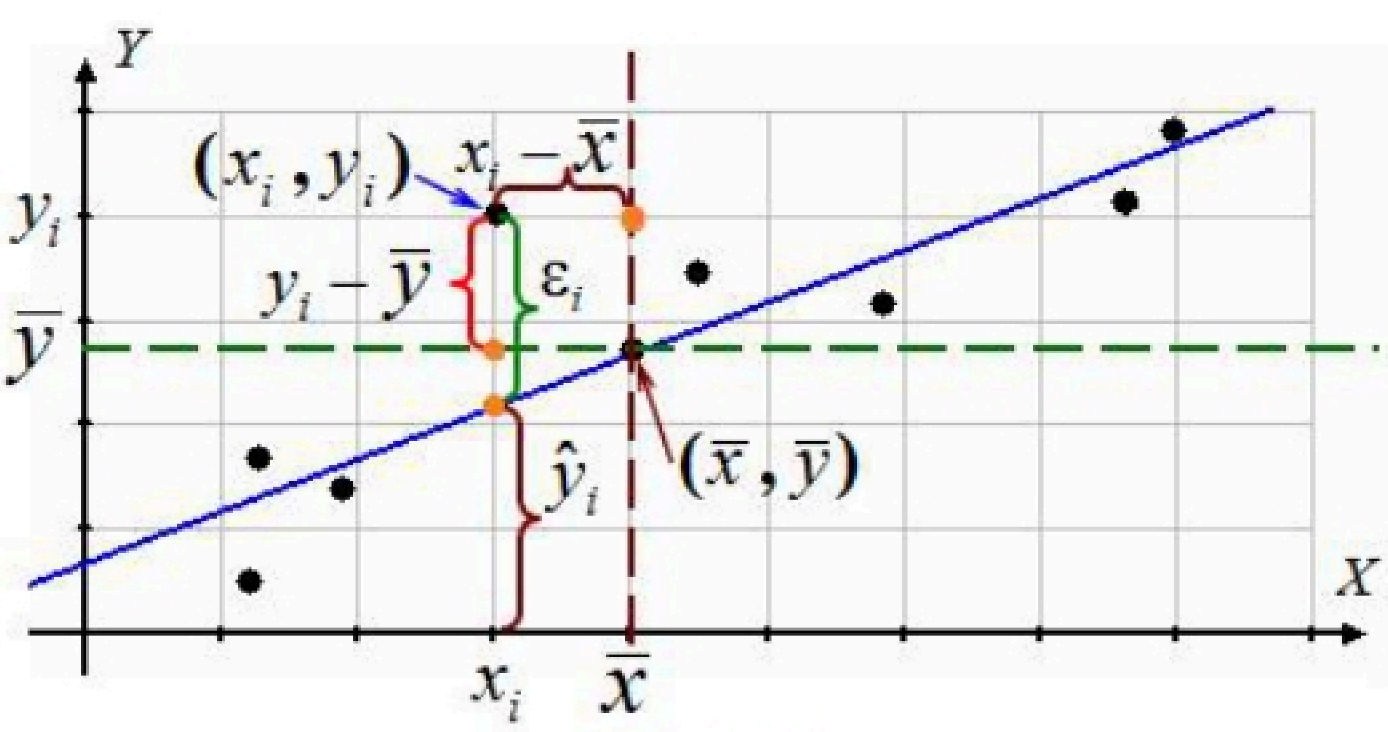
٢- من أجل مستقيم انحدار $Y$ على $X$، يُدعى المقدار:

**التغيُّر الكلي** Total Variation للمتغيِّر $Y$
$$\mathbf{TV}_Y := \sum_{i=1}^{n}\left(y_i-\overline{y}\right)^2$$

**التغيُّر المُفسَّر** Explained Variation للمتغيِّر $Y$
$$\mathbf{EV}_Y := \sum_{i=1}^{n}\left(\hat{y}_i-\overline{y}\right)^2$$

**التغيُّر غير المُفسَّر** Unexplained Variation للمتغيِّر $Y$
$$\mathbf{UV}_Y := \sum_{i=1}^{n}\left(y_i-\hat{y}\right)^2$$

كما هو واضح من هذه التسميات، فإنَّ التغيُّر المُفسَّر يمكن أن يُفسَّر بوساطة علاقة الانحدار (العلاقة بين $X$ و$Y$)، بمعنى أنَّ التغيُّر المُفسَّر هو الجزء من التغيُّر الكلي لـ $Y$ الذي يمكن إرجاعه إلى التغيُّر في المتغيِّر المستقل $X$. أمّا التغيُّر غير المُفسَّر فإنَّه لا يمكن تفسيره بوساطة علاقة الانحدار وإنَّما هو نتيجة للمصادفة أو بسبب متغيِّرات أخرى، وذلك لأنَّ هذا الاختلاف يُمثِّل مجموع مربّعات المسافات الرأسية التي تمثِّل انحرافات القيم الحقيقية (المشاهدة) عن خط الانحدار. إذاً، فهو يقيس الجزء من التغيُّر الكلي الذي يعود إلى عوامل أخرى غير المتغيِّر المستقل $X$. أي إن التغيُّر في المتغيِّر المستقل $X$ لا يُفسِّر هذا الجزء من التغيُّر الكلي (انظر الشكل التوضيحي الآتي).




شكل (٣,٢٣)

وهكذا يكون مجموع التغيُّر المُفسَّر وغير المُفسَّر يساوي التغيُّر الكلي، أي إنَّ $\mathbf{TV}_Y = \mathbf{EV}_Y + \mathbf{UV}_Y$، كما يمكن حساب قيمة مُعامل التحديد وفقاً للعلاقة الآتية:

$$\mathbf{r}^2 = \frac{\mathbf{EV}_Y}{\mathbf{TV}_Y} = \frac{\sum_{i=1}^{n}\left(\hat{y}_i - \overline{y}\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2} \qquad [3,28]$$

وبالمثل يكون لدينا من أجل مستقيم انحدار $X$ على $Y$ ما يلي:

التغيُّر الكلي للمتغيِّر $X$ هو $$\mathbf{TV}_X = \sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2$$

التغيُّر المُفسَّر للمتغيِّر $X$ هو $$\mathbf{EV}_X = \sum_{i=1}^{n}\left(\hat{x}_i - \overline{x}\right)^2$$

التغيُّر غير المُفسَّر للمتغيِّر $X$ هو $$\mathbf{UV}_X = \sum_{i=1}^{n}\left(x_i - \hat{x}\right)^2$$

٣- إذا وقعت كل المشاهدات على خط الانحدار فعندئذ سيكون $\mathbf{UV}_Y = 0$، ومن ثَمَّ يصبح $\mathbf{TV}_Y = \mathbf{EV}_Y$، وهذا يعني أنَّ كل التغيُّر في $Y$ يرجع إلى التغيُّر في المتغيِّر المستقل $X$. بعبارة أخرى، إنَّ التغيُّر في المتغيِّر المستقل $X$ يشرح كل التغيُّر في $Y$، وبما أنَّ قيمة مُعامل الارتباط في هذه الحالة سوف يساوي الواحد تماماً، فإنَّ قيمة مُعامل التحديد ستصبح مساويةً الواحد أيضاً، وهذا يعني أنَّه يمكن استخدام قيمة مُعامل التحديد للحكم على جودة التوفيق لمستقيم الانحدار، فكلّما كانت قيمة هذا المُعامل أقرب إلى الواحد؛ دلَّ ذلك على جودة أفضل لتوفيق البيانات بمستقيم الانحدار، والعكس بالعكس.

### (٣,٢,٤,٣) أمثلة

رصدت الزيادة في درجات الحرارة لآلة بعد مرور أكثر من ساعة من بدء عملها وعلى فترات زمنية متتالية، فكانت النتائج كما في الجدول الآتي:

الجدول (٢٢, ٣. أ)

| $i$ | $x_i$ | $y_i$ | $i$ | $x_i$ | $y_i$ | $i$ | $x_i$ | $y_i$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.07 | 8.435 | 9 | 2.5 | 9.59 | 17 | 9.9 | 12.007 |
| 2 | 1.19 | 8.612 | 10 | 2.9 | 9.82 | 18 | 11.4 | 12.331 |
| 3 | 1.33 | 8.743 | 11 | 3.4 | 10.035 | 19 | 12.9 | 12.605 |
| 4 | 1.47 | 8.943 | 12 | 4.1 | 10.318 | 20 | 14.4 | 12.89 |
| 5 | 1.63 | 8.986 | 13 | 4.9 | 10.596 | 21 | 16.4 | 13.147 |
| 6 | 1.8 | 9.153 | 14 | 5.9 | 10.879 | 22 | 18.9 | 13.385 |
| 7 | 2 | 9.274 | 15 | 7.1 | 11.26 | 23 | 21.9 | 13.577 |
| 8 | 2.2 | 9.415 | 16 | 8.4 | 11.645 | | | |

علماً أنَّ $x_i$ هي اللحظات الزمنية التي أخذت عندها القياســات و $y_i$ هي قيم الزيادة في درجة الحرارة عند اللحظة $x_i$، ولنقم بتعيين معادلة مستقيم المربعات الصغرى لـ $Y$ على $X$ مستعينين بالجدول الآتي (١٩, ٣. ب).

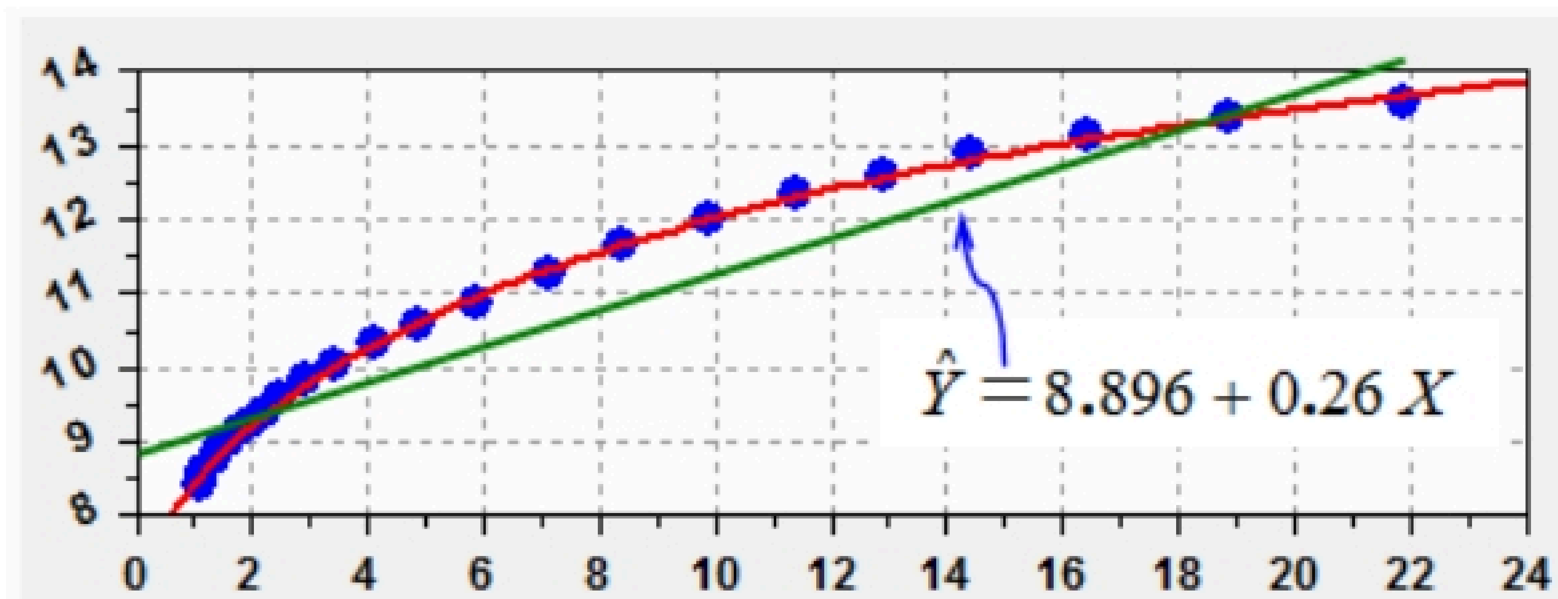
الجدول (٢٢, ٣. ب)

| $i$ | $x_i$ | $x_i^2$ | $y_i$ | $y_i^2$ | $x_i \cdot y_i$ |
|---|---|---|---|---|---|
| 1 | 1.07 | 1.1449 | 8.435 | 71.14923 | 9.02545 |
| 2 | 1.19 | 1.4161 | 8.612 | 74.16654 | 10.24828 |
| 3 | 1.33 | 1.7689 | 8.743 | 76.44005 | 11.62819 |
| 4 | 1.47 | 2.1609 | 8.943 | 79.97725 | 13.14621 |
| 5 | 1.63 | 2.6569 | 8.986 | 80.7482 | 14.64718 |
| 6 | 1.8 | 3.24 | 9.153 | 83.77741 | 16.4754 |
| 7 | 2 | 4 | 9.274 | 86.00708 | 18.548 |
| 8 | 2.2 | 4.84 | 9.415 | 88.64223 | 20.713 |
| 9 | 2.5 | 6.25 | 9.59 | 91.9681 | 23.975 |
| 10 | 2.9 | 8.41 | 9.82 | 96.4324 | 28.478 |
| 11 | 3.4 | 11.56 | 10.035 | 100.7012 | 34.119 |
| 12 | 4.1 | 16.81 | 10.318 | 106.4611 | 42.3038 |
| 13 | 4.9 | 24.01 | 10.596 | 112.2752 | 51.9204 |
| 14 | 5.9 | 34.81 | 10.879 | 118.3526 | 64.1861 |
| 15 | 7.1 | 50.41 | 11.26 | 126.7876 | 79.946 |
| 16 | 8.4 | 70.56 | 11.645 | 135.606 | 97.818 |
| 17 | 9.9 | 98.01 | 12.007 | 144.168 | 118.8693 |
| 18 | 11.4 | 129.96 | 12.331 | 152.0536 | 140.5734 |
| 19 | 12.9 | 166.41 | 12.605 | 158.886 | 162.6045 |
| 20 | 14.4 | 207.36 | 12.89 | 166.1521 | 185.616 |
| 21 | 16.4 | 268.96 | 13.147 | 172.8436 | 215.6108 |
| 22 | 18.9 | 357.21 | 13.385 | 179.1582 | 252.9765 |
| 23 | 21.9 | 479.61 | 13.577 | 184.3349 | 297.3363 |
| sum | 157.69 | 1951.568 | 245.646 | 2687.089 | 1910.765 |

فنجد أنَّ قيم المَعْلمتين $a$ و $b$ هما $a = 8.896$ و $b = 0.26$، ومن ثمَّ يكون لمعادلة مستقيم انحدار $Y$ على $X$ العلاقة الآتية:

$$\hat{Y} = 8.896 + 0.26X$$

ولتمثيلها البياني الشكل الآتي:




الشكل (٣,٢٤)

ومن أجل حساب الخطأ المعّياري للتقدير فإنّه لدينا:

الجدول (٣,٢٢.ج)

| $i$ | $x_i$ | $y_i$ | $\varepsilon_i$ | $\hat{y}_i$ | $(y_i-\bar{y})^2$ | $(\hat{y}_i-\bar{y})^2$ | $(y_i-\hat{y}_i)^2$ |
|---|---|---|---|---|---|---|---|
| 1 | 1.07 | 8.435 | -0.739 | 9.174002 | 5.040025 | 2.268031 | 0.546123 |
| 2 | 1.19 | 8.612 | -0.59324 | 9.20524 | 4.276624 | 2.174916 | 0.351934 |
| 3 | 1.33 | 8.743 | -0.49869 | 9.241686 | 3.751969 | 2.068747 | 0.248688 |
| 4 | 1.47 | 8.943 | -0.33513 | 9.278131 | 3.017169 | 1.965236 | 0.112313 |
| 5 | 1.63 | 8.986 | -0.33378 | 9.319783 | 2.869636 | 1.85019 | 0.111411 |
| 6 | 1.8 | 9.153 | -0.21104 | 9.364038 | 2.331729 | 1.731755 | 0.044537 |
| 7 | 2 | 9.274 | -0.1421 | 9.416103 | 1.976836 | 1.597435 | 0.020193 |
| 8 | 2.2 | 9.415 | -0.05317 | 9.468168 | 1.600225 | 1.468537 | 0.002827 |
| 9 | 2.5 | 9.59 | 0.043735 | 9.546265 | 1.1881 | 1.285354 | 0.001913 |
| 10 | 2.9 | 9.82 | 0.169605 | 9.650395 | 0.7396 | 1.060086 | 0.028766 |
| 11 | 3.4 | 10.035 | 0.254443 | 9.780557 | 0.416025 | 0.808997 | 0.064741 |
| 12 | 4.1 | 10.318 | 0.355216 | 9.962784 | 0.131044 | 0.514398 | 0.126178 |
| 13 | 4.9 | 10.596 | 0.424956 | 10.17104 | 0.007056 | 0.259036 | 0.180588 |
| 14 | 5.9 | 10.879 | 0.447632 | 10.43137 | 0.039601 | 0.061818 | 0.200374 |
| 15 | 7.1 | 11.26 | 0.516243 | 10.74376 | 0.3364 | 0.004065 | 0.266506 |
| 16 | 8.4 | 11.645 | 0.562821 | 11.08218 | 0.931225 | 0.161748 | 0.316767 |
| 17 | 9.9 | 12.007 | 0.534334 | 11.47267 | 1.760929 | 0.628319 | 0.285513 |
| 18 | 11.4 | 12.331 | 0.467848 | 11.86315 | 2.725801 | 1.399849 | 0.218882 |
| 19 | 12.9 | 12.605 | 0.351361 | 12.25364 | 3.705625 | 2.476338 | 0.123455 |
| 20 | 14.4 | 12.89 | 0.245875 | 12.64413 | 4.8841 | 3.857787 | 0.060454 |
| 21 | 16.4 | 13.147 | -0.01777 | 13.16477 | 6.086089 | 6.174101 | 0.000316 |
| 22 | 18.9 | 13.385 | -0.43058 | 13.81558 | 7.317025 | 9.831891 | 0.185403 |
| 23 | 21.9 | 13.577 | -1.01956 | 14.59656 | 8.392609 | 15.33942 | 1.039498 |
| sum | 157.69 | 245.646 | 5.4E-10 | 245.646 | 63.52544 | 58.98806 | 4.537382 |

حيث نجد من مُعطيات الجدول السّابق ما يلي:

- التغيُّر الكلي للمتغيِّر $Y$ يساوي

$$\mathbf{TV}_Y = \sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2 = 63.52544$$

- التغيُّر المُفسَّر للمتغيِّر $Y$ يساوي

$$\mathbf{EV}_Y = \sum_{i=1}^{n}\left(\hat{y}_i - \overline{y}\right)^2 = 58.98806$$

- التغيُّر غير المُفسَّر للمتغيِّر $Y$ يساوي

$$\mathbf{UV}_Y = \sum_{i=1}^{n}\left(y_i - \hat{y}\right)^2 = 4.537382$$

وهكذا يكون لدينا:

$$\mathbf{EV}_Y + \mathbf{UV}_Y = 58.98806 + 4.537382 = 63.52544 = \mathbf{TV}_Y$$

ومن جهة أخرى نجد أنَّ الخطأ المعياري لتقدير مستقيم الانحدار يساوي:

$$S_{est\hat{Y}} = \sqrt{\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n-2}} = \sqrt{\frac{4.53738}{23-2}} = 0.46483$$

وبحساب قيمة الانحراف المعياري $S_Y$ نجد $S_Y = 1.69927$، ومن ثَمَّ باستخدام العلاقة [3,26] نجد أنَّ قيمة مُعامل التحديد $\mathbf{r}^2$ تساوي:

$$\mathbf{r}^2 = 1 - \frac{(n-2)S^2_{est\hat{Y}}}{(n-1)S^2_Y} = 1 - \frac{21 \times 0.2161}{22 \times 2.88752} = 0.92857$$

أو باستخدام العلاقة [3,28] يكون لدينا:

$$\mathbf{r}^2 = \frac{\mathbf{EV}_Y}{\mathbf{TV}_Y} = \frac{\sum_{i=1}^{n}\left(\hat{y}_i - \overline{y}\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}\right)^2} = \frac{58.98806}{63.52544} = 0.92857$$

إنَّ العلاقة الأخيرة توضِّح لنا أنَّ 93% من أسباب الزيادة في ارتفاع الحرارة تعود لطول فترة تشغيل الآلة، وأكثر من ذلك نلاحظ أنَّ $\mathbf{r} = 0.9636$ التي تثبت وجود ارتباط قوي جداً بين ازدياد درجة الحرارة وطول الفترة الزمنية التي تعمل بها الآلة.

الآن، وبإمعان النظر في الشكل (٣,٢٤) فإنَّنا نلاحظ أنَّ خط الانحدار الأنسب للبيانات هو خط منحنٍ وليس مستقيماً، ومعادلة ذلك المنحنى تُدعى نموذج ويبول Weibull Model للانحدار (نسبة إلى المهندس والرياضياتي السويدي ويبول Ernst Hjalmar Waloddi Weibull (1887-1979))، ومعادلته من الشكل:

$$\hat{Y} = 16269 - 9.5658\,\mathbf{exp}\left[-20.3712\,X^{60.1963}\right]$$

وبخطأ معياري يساوي $S_{est\hat{Y}} = 0.0602$، وهذا يعني أنَّ توفيق هذه البيانات بهذه المعادلة ممتازاً، و أمّا مُعَامل التحديد الموافق لهذا المنحنى فإنَّه يساوي $\mathbf{r}^2 = 0.998$ الذي يوضِّح لنا أنَّ 99.89% من أسباب الزيادة في ارتفاع حرارة الآلة تعود لطول فترة تشغيل هذه الآلة.

**لكن لماذا تناولنا هذا الحوار وقد أكدنا سابقاً بأنَّنا سنهتمّ بالنماذج الخطِّية فقط؟**

في الحقيقة توجد حالات يكون فيها الحصول على منحنى لتوفيق انحدار البيانات مُمكناً من خلال استخدام تقنية الانحدار الخطي، ومن ثَمَّ نحصل على خط انحدار أفضل (أي بخطأ معيّاري أقل) وهو بدوره يقدّم لنا تقديرات أكثر دقّة من مستقيم الانحدار، والفقرة التالية تقدّم لنا عرضاً موجزاً حول هذا الموضوع.

## (٣,٣) الانحدار غير الخطي وتوفيق المنحنيات

### Nonlinear Regression and Curves Fitting

لقد لاحظنا في أمثلة سابقة أنَّه قد تكون طبيعة العلاقة بين المتغيِّرين $X$ و$Y$ غير خطيّة، ومن ثمَّ يكون استخدام مستقيم الانحدار لتقدير قيم المشاهدات غير مجُسبب كِبَر الخطأ المعياري الذي سيرافقه، وذلك لأنَّه قد يكون قسم كبير من النقاط الممثِّلة للمشاهدات متناثرة بعيداً نسبياً عن مستقيم الانحدار (انظر الشكل التوضيحي (٢١, ٣ .ب))، ولذلك فإنَّ محاولة تمهيد الخط البياني (إن أمكن ذلك) الممثِّل للعلاقة المتوسطة بين المتغيِّرين سيكون أكثر نفعاً وفائدة، ومن ثَمّ تصبح الفروق ما بين القيم المشاهدة والقيم المقدَّرة لها صغيرةً نسبياً، وهذا بدوره يجعل قيمة مُعامل الارتباط (ومن ثمَّ قيمة مُعامل التحديد) أكبر بكثير من تلك التي لمستقيم الانحدار. لهذا السبب كنَّا منذ البدء حذرين في تفسير النتائج المترتّبة على قيمة مُعامل الارتباط.

الآن لنفترض أنَّ لدينا بيانات على شكل ثنائيات (أزواج) ناتجة عن دراسة متغيِّرين $X$ و$Y$ (ممثِّلين لظاهرتين)، فعندئذ من أجل بعض الحالات الخاصة يمكننا ردّ العلاقات غير الخطّية للانحدار إلى شكل خطي من خلال استخدام تحويلات مناسبة للبيانات، ومن ثمَّ دراسة البيانات الناتجة عن التحويل وفقاً للانحدار الخطي، وبعد ذلك العودة إلى العلاقة الأصل باستخدام التحويلات العكسية للعلاقات التي استخدمت في البدء. لكن قبل البدء بتقديم بعض هذه الحالات الخاصّة سنقوم بتوضيح عملية الخطّية لعلاقة (ممكنة التحويل) وذلك باستخدام المثال السّابق (٣, ٤, ٢, ٣) حيث وجدنا أنَّ البيانات المُعطاة تمتلك انحداراً من نموذج ويبول، ومعادلته كانت من الشكل الآتي:

$$Y = 16269 - 9.5658\,\mathbf{exp}\left[-20.3712\,X^{60.1963}\right] \qquad (*)$$

ولنشاهد ما الذي سيحدث للبيانات بعد استخدام التحويلات اللازمة لعمليّة الخطّية.

بملاحظ أنَّه يمكننا كتابة العلاقة السابقة (*) على النحو الآتي:

$$\mathbf{ln}\left[\frac{Y-16269}{-9.5658}\right] = -20.3712\,X^{60.1963}$$

والتي تُكتب بالشكل الآتي أيضاً:

$$\mathbf{ln\,ln}\left[\frac{Y-16269}{-9.5658}\right]^{-1} = \mathbf{ln}\,20.3712 + 60.1963\ \ \mathbf{ln}\,X$$

وينتج عنها أنّه من أجل كل $i \in \mathbb{N}_{23}$ سيكون لدينا:

$$\mathbf{ln\,ln}\left[\frac{y_i-16269}{-9.5658}\right]^{-1} = \mathbf{ln}\,20.3712 + 60.1963\ \ \mathbf{ln}\,x_i$$

والآن لنضع $b = 60.1963$، ولنستخدم التحويلات:

$$Y^* := \mathbf{ln\,ln}\left[\frac{Y-16269}{-9.5658}\right]^{-1}$$

$$a^* := \mathbf{ln}\,20.3712$$

$$X^* := \mathbf{ln}\,X$$

فيصبح للعلاقة السابقة (*) العرض:

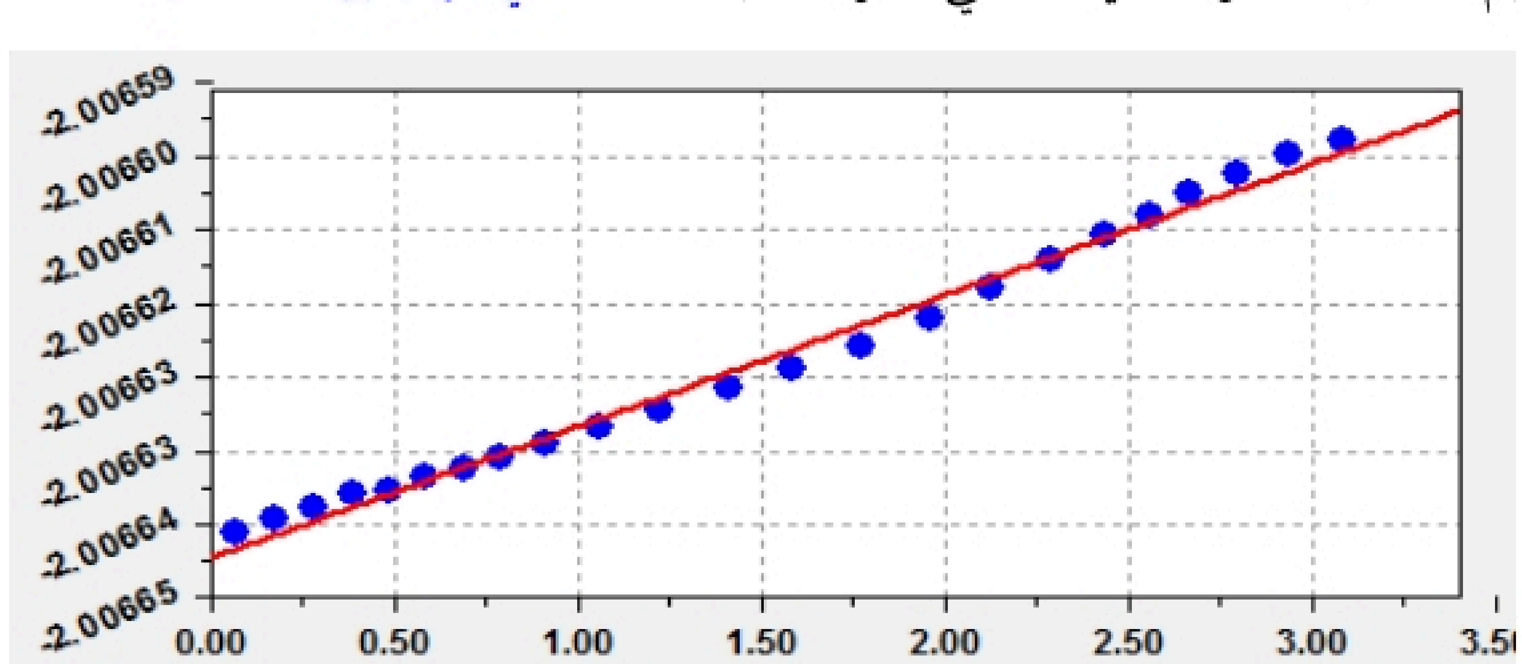
$$Y^* = a^* + bX^*$$

وهي علاقة خطيّة بـ $X^*$ ، والجدول الآتي يدوِّن قيم $x_i^*$ و $y_i^*$ .

الجدول (٢٢,٣.د)

| $i$ | $x_i$ | $y_i$ | $x_i^*$ | $y_i^*$ |
|---|---|---|---|---|
| 1 | 1.07 | 8.435 | 0.068 | -2.00664285 |
| 2 | 1.19 | 8.612 | 0.174 | -2.00664139 |
| 3 | 1.33 | 8.743 | 0.285 | -2.0066403 |
| 4 | 1.47 | 8.943 | 0.385 | -2.00663865 |
| 5 | 1.63 | 8.986 | 0.489 | -2.0066383 |
| 6 | 1.8 | 9.153 | 0.588 | -2.00663692 |
| 7 | 2 | 9.274 | 0.693 | -2.00663591 |
| 8 | 2.2 | 9.415 | 0.788 | -2.00663475 |
| 9 | 2.5 | 9.59 | 0.916 | -2.0066333 |
| 10 | 2.9 | 9.82 | 1.065 | -2.0066314 |
| 11 | 3.4 | 10.035 | 1.224 | -2.00662962 |
| 12 | 4.1 | 10.318 | 1.411 | -2.00662728 |
| 13 | 4.9 | 10.596 | 1.589 | -2.00662498 |
| 14 | 5.9 | 10.879 | 1.775 | -2.00662264 |
| 15 | 7.1 | 11.26 | 1.96 | -2.00661949 |
| 16 | 8.4 | 11.645 | 2.128 | -2.00661631 |
| 17 | 9.9 | 12.007 | 2.293 | -2.00661332 |
| 18 | 11.4 | 12.331 | 2.434 | -2.00661064 |
| 19 | 12.9 | 12.605 | 2.557 | -2.00660837 |
| 20 | 14.4 | 12.89 | 2.667 | -2.00660601 |
| 21 | 16.4 | 13.147 | 2.797 | -2.00660389 |
| 22 | 18.9 | 13.385 | 2.939 | -2.00660192 |
| 23 | 21.9 | 13.577 | 3.086 | -2.00660033 |

والآن لو نظرنا إلى العرض الانتشاري لهذه النقاط $(x_i^*, y_i^*)$ مع $i$ و $\mathrm{N}_{23}$ ، فإنّنا سنلاحظ كيف أصبح شكل توزّع البيانات وكأنّها واقعة على خط مستقيم واحد،وهذا هو التفسير الهندسي لمدلول خطيّة علاقة ما (أي تحويلها إلى علاقة خطيّة).




شكل (٣,٢٥)

### (٣,٣,١) نماذج انحدار يمكن ردّها إلى خطيّة

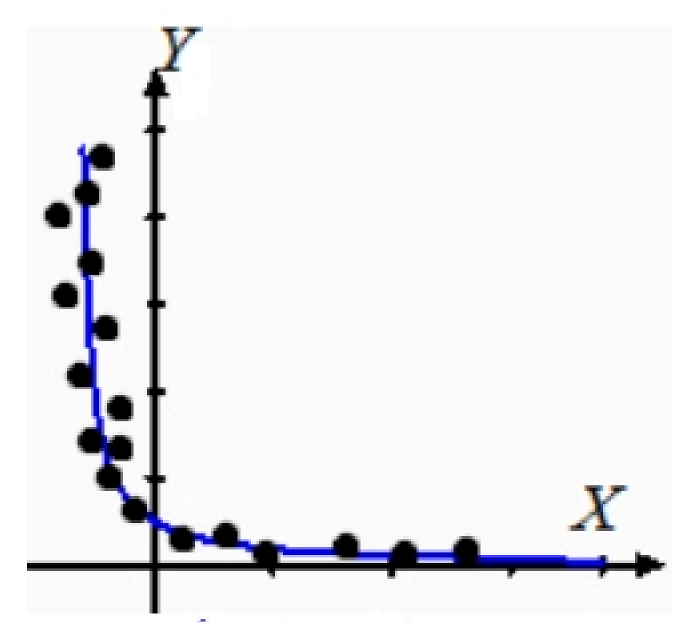
فيما يلي نقدّم بعض العلاقات البسيطة والشهيرة التي يمكن خطيّتها، ونبدؤها بالانحدار الكسري البسيط (أو التناظري):

#### (٣,٣,١,١) الانحدار الكسري البسيط Simple Fractional Regression

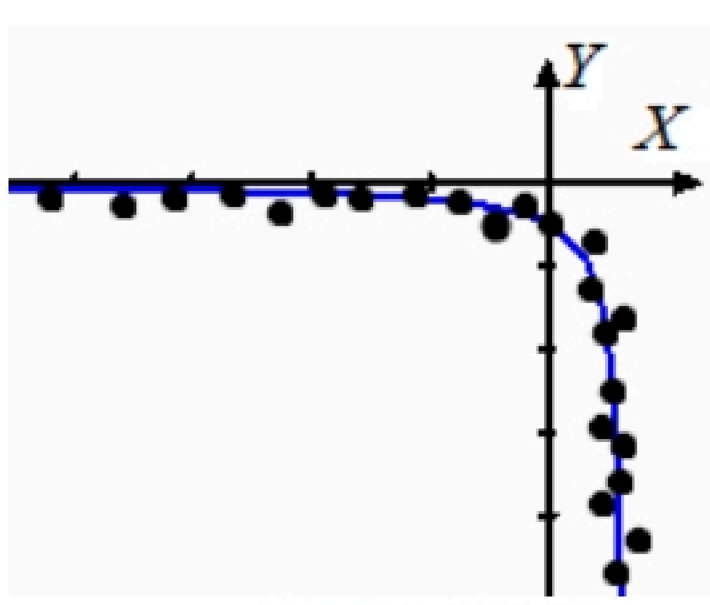
في هذا النوع من الانحدار تكون العلاقة التي تربط بين المتغيِّرين $X$ و$Y$ من الشكل الآتي:

$$Y = \frac{1}{a+bX} \qquad ; a, b \in \mathbb{R}, b \neq 0 \tag{3,30}$$

نلاحظ هنا أنَّ الدالَّة المُعطاة بالعلاقة السابقة هي علاقة دالَّة كسرية بسيطة (أو تناظرية)، ولذلك فإنَّ الانحدار الذي يستخدم هذه العلاقة يُسمّى انحدارًا كسريًّا بسيطًا (انظر الأشكال التوضيحية الآتية).



الشكل (٣,٢٦.أ)



الشكل (٣,٢٦.ب)

فعندئذ من أجل خطيّة هذه العلاقة نقوم باستخدام التحويل:

$$y_i^* = \frac{1}{y_i} \qquad ; i = 1,2,\dots,n$$

ومن ثمَّ يكون لعلاقة الانحدار الشكل الخطي الآتي:

$$Y^* = a + bX \qquad ; a, b \in \mathbb{R}, b \neq 0$$

وبعد ذلك نقوم بحساب الثوابت $b$ و$a$ بطريقة المربعات الصغرى بناءً على البيانات بعد عملية التحويل، أي يجب الأخذ بالحسبان أنَّ القيم لـ $Y^*$ التي سنعوض بها في العلاقات لتعيين $a$ و $b$ هي $y_i^* = \frac{1}{y_i}$ مع $i \in \mathbb{N}_n$، ومن ثمَّ التعويض في العلاقة السابقة، وبعد ذلك العودة إلى العلاقة الأصل من خلال وضع:

$$\hat{Y} = \frac{1}{Y^*} = \frac{1}{a+bX} \qquad ; a, b \in \mathbb{R}, b \neq 0$$

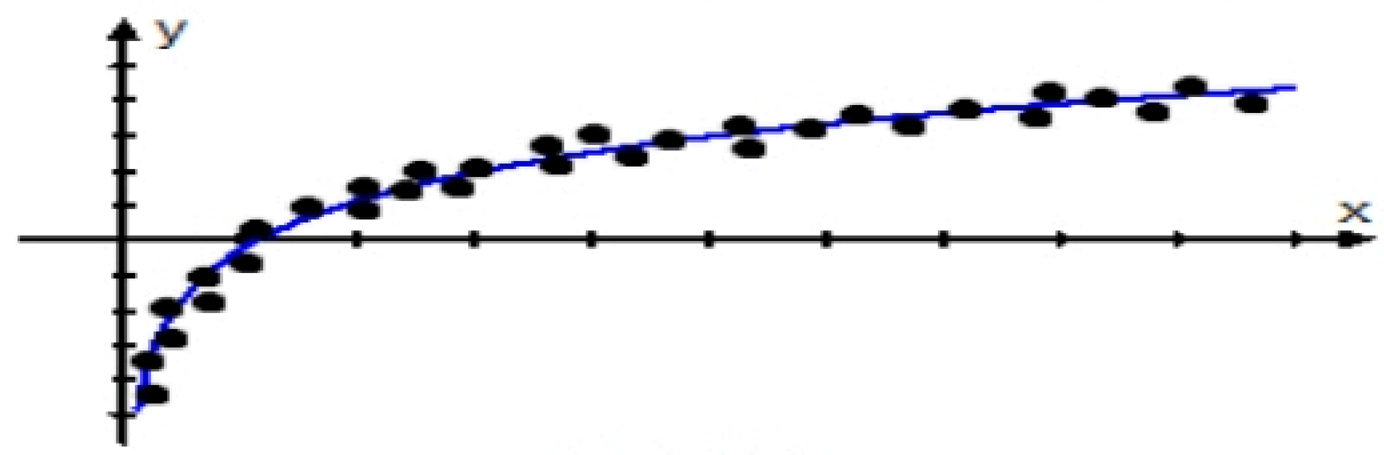
فنكون بذلك قد عينًّا علاقة منحنى الانحدار.

#### (٣,٣,١,٢) الانحدار اللوغاريتمي Logarithmic Regression

في هذا النوع من الانحدار تكون العلاقة التي تربط بين المتغيِّرين $X$ و$Y$ من الشكل الآتي:

$$Y = a + b\ln X \qquad ; X > 0 \ \& \ a, b \in \mathbb{R}, b \neq 0 \tag{3,31}$$

لاحظ أنَّ الدالَّة المُعطاة بالعلاقة السابقة هي علاقة دالَّة لوغاريتميّة، ولذلك فإنَّ الانحدار الذي يستخدم هذه العلاقة يُسمّي **الانحدار اللوغاريتمي** (انظر الشكل الآتي لبيانات تتوزَّع وفقاً للعلاقة [3,31]).




الشكل (٣,٢٧)

وفي هذه الحالة نستخدم التحويل $X^* = \ln X$ فتصبح لدينا العلاقة:

$$Y = a + b X^* \quad ; a , b \in \mathbb{R} , b \neq 0$$

وبعد ذلك نقوم بتعيين قيم الثوابت $a$ و $b$ بطريقة المربّعات الصغرى على أنَّ نأخذ بالحسبان أنَّ القيم $X^*$ التي سنعوض بها في العلاقات لتعيين $a$ و $b$ هي $\ln x_i$ لكلّ $i \in \mathbb{N}_n$، وبعد ذلك نعوض لنحصل على العلاقة:

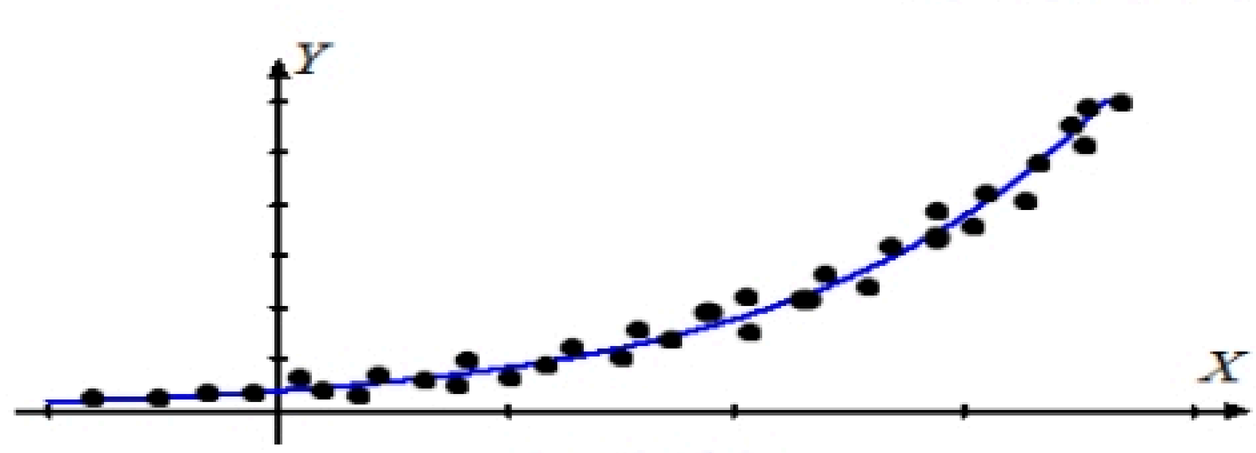
$$\hat{Y} = a + b \ln X \quad ; X > 0 \ \& \ a , b \in \mathbb{R} , b \neq 0$$

### (٣,٣,١,٣) الانحدار الأسّي Exponential Regression

في هذا النوع من الانحدار تكون العلاقة التي تربط بين المتغيِّرين $X$ و $Y$ من الشكل الآتي:

$$Y = a\,\mathbf{e}^{bX} \quad ; \ a , b \in \mathbb{R}, \ b \neq 0 \qquad [3,32]$$

فلاحظ أنَّ الدالَّة المُعطاة بالعلاقة السابقة هي علاقة لدالَّة أسّيّة، ولذلك فإنَّ الانحدار الذي يستخدم هذه العلاقة يُسمّي **الانحدار الأسّي** (انظر الشكل الآتي لبيانات تتوزَّع وفقاً للعلاقة [3,32]).




الشكل (٣,٢٨)

في هذه الحالة نقوم بردّ هذا الشكل إلى شكل خطي من خلال أخذ اللوغاريتم لطرفي العلاقة السابقة فيصبح لها العرض الآتي:

$$\ln Y = \ln a + b X$$

ومن ثَّم نستخدم التحويلين $a^* = \ln a$ و $Y^* = \ln Y$، فيكون لدينا علاقة خطّية من الشكل:

$$Y^* = a^* + b X$$

وبعد ذلك نقوم بتعيين قيم الثوابت $a^*$ و $b$ بوساطة طريقة المربّعات الصغرى بناءً على البيانات بعد عملية التحويل، ومن ثَّم نقوم بالتعويض في العلاقة الأخيرة، وبعد ذلك العودة إلى الشكل الأصل باستخدام العلاقات العكسية فنحصل على تقدير $Y$ من خلال العلاقة:

$$\hat{Y} = \mathbf{e}^{Y^*} = \mathbf{e}^{a^* + bX} = \mathbf{e}^{\ln a}\ \mathbf{e}^{bX} = a\ \mathbf{e}^{bX}$$

### (٣,٣,١,٤) انحدار القوى Power Regression

إذا كانت علاقة الانحدار التي تربط بين المتغيِّرين $X$ و $Y$ من الشكل:

$$Y = aX^{b} \quad ; X > 0 \ \& \ a > 0 \ , \ b \in \mathbb{Q} \qquad [3,33]$$

وهذا النوع من الانحدار يُدعى **انحدار القوى** لأنَّ الدالَّة المُعطاة بالعلاقة السابقة هي علاقة لدالَّة قوى بـ $X$ (**انظر الأشكال الآتية**) التي يمكن تحويلها إلى نموذج خطي من خلال أخذ اللوغاريتم لطرفي العلاقة السابقة، فنحصل على العلاقة:

$$\ln Y = \ln a + b \ln X$$

ومن ثَّم نستخدم التحويلات:

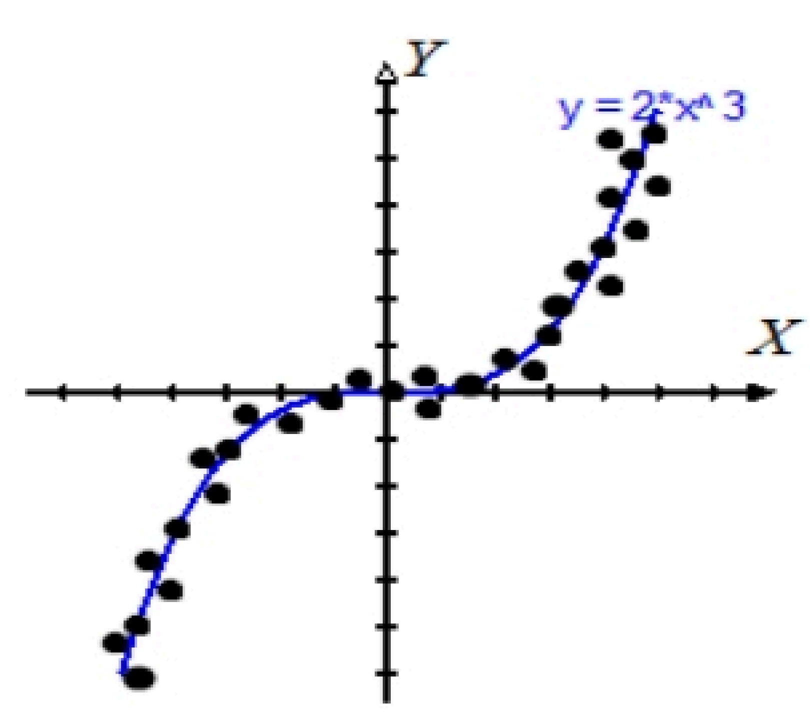
$$Y^{*} = \ln Y \qquad \& \qquad a^{*} = \ln a \qquad \& \qquad X^{*} = \ln X$$

فيصبح لدينا:

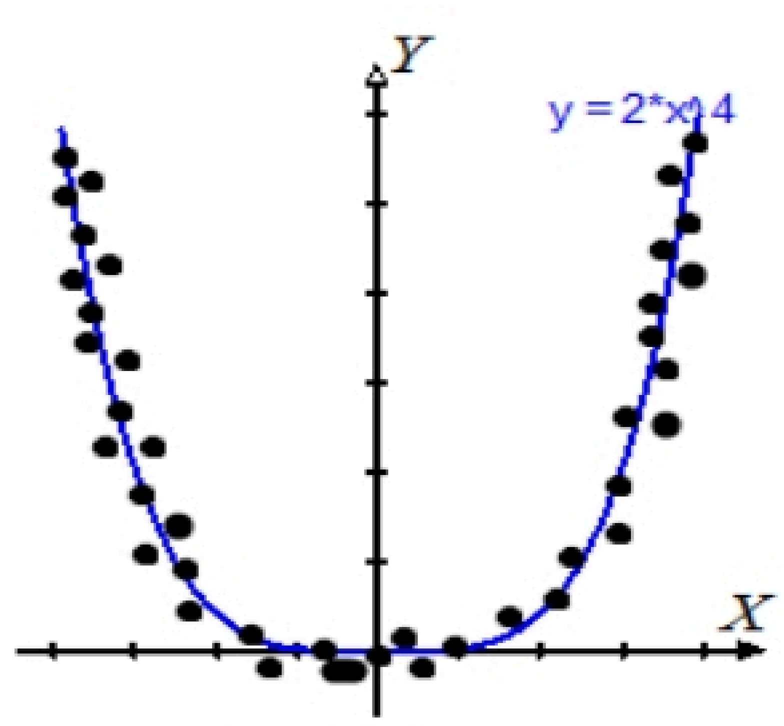
$$Y^{*} = a^{*} + b X^{*}$$

وبعد ذلك نقوم بتقدير $a^{*}$ و $b$ بطريقة المربّعات الصغرى، ومن ثَّم التعويض في العلاقة الأخيرة بناءً على البيانات بعد عملية التحويل، ومن ثَّم التعويض في العلاقة السابقة، وبعد ذلك العودة إلى الشكل الأصل باستخدام العلاقات العكسية فنحصل على تقدير من خلال:

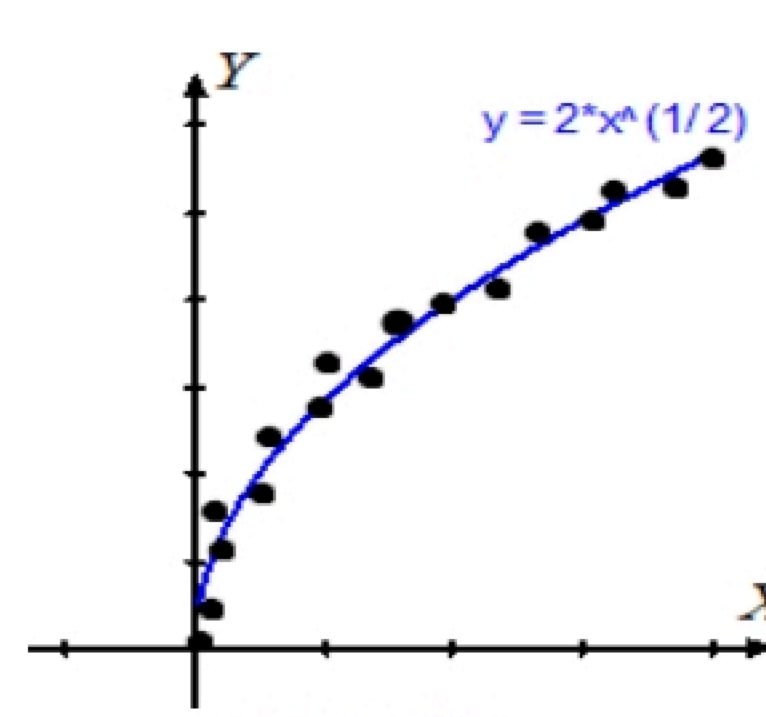
$$\hat{Y} = e^{Y^{*}} = e^{a^{*} + bX^{*}} = e^{a^{*}} e^{bX^{*}} = e^{\ln a} e^{b \ln X} = aX^{b}$$



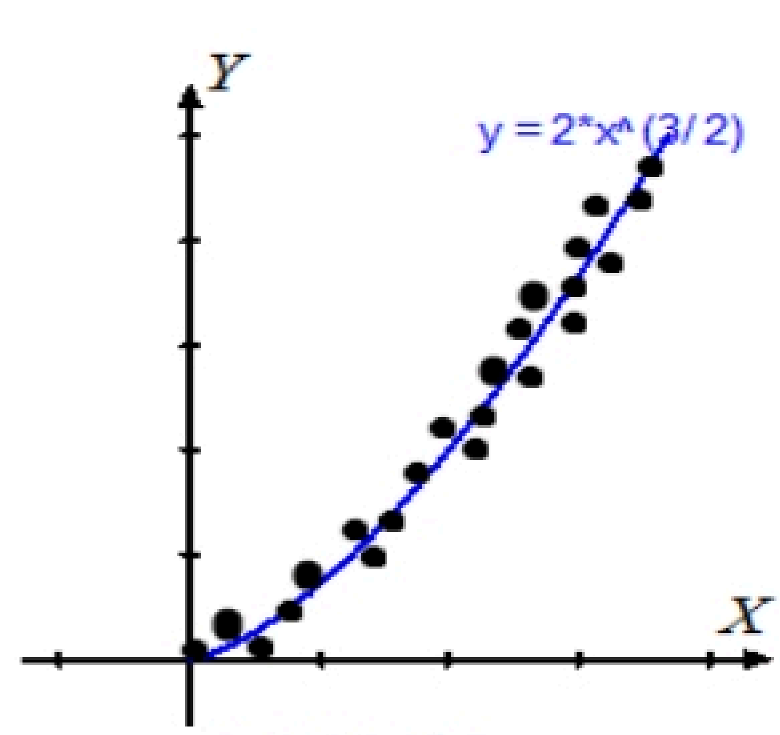

الشكل (٣,٢٩.أ)




الشكل (٣,٢٩.ب)




الشكل (٣,٣٠.ج)




الشكل (٣,٣٠.د)

### (٣,٣,٢) نماذج لا يمكن خطيتها

لقد ذكرنا فيما سبق بأنَّه يمكن ردّ بعض علاقات الانحدار غير الخطيّة إلى نماذج خطيّة، ومن ثمَّ دراستها كما هو الحالة لدى الانحدار الخطي لتقدير الثوابت $a$ و$b$ في معادلة الانحدار، ولكن هناك علاقات كثيرة لا يمكن ردّها إلى شكل خطي منها على سبيل المثال لا الحصر كثيرات الحدود $Y = a_o + a_1 X + a_2 X^2 + \ldots + a_n X^n$ مع $n \in \mathbb{N}$، وكذلك العلاقات المثلثيّة التي من الشكل:

$$Y = a \cos bX \quad \text{or} \quad Y = a \sin bX$$

وهناك علاقات كثيرة أخرى. في مثل هذه الحالات يمكن معالجة بعض الحالات الخاصة منها بطريقة المربّعات الصغرى، أو بطرائق أخرى مثل طريقة الفروق المتتالية، ولكنَّنا لن نتطرق لمثل هذه الطرائق في هذا الفصل حيث سنقوم في فصل قادم (الفصل الثاني عشر) باستخدام طريقة الفروق المتتالية لتوفيق منحنى لكثيرة حدود من الدرجة $n$ من أجل استخدامه كتقدير للاتجاه العام لمتسلسلة زمنية.

## (٣,٤) الارتباط المتعد والجزئي

## Multiple and Partial Correlation

لقد لاحظنا أنَّه عندما يكون لدينا متغيِّرين يمكن البحث في العلاقة التي تربط بينهما من خلال حساب تغايرهما كما في مُعامل بيرسون للارتباط الخطي، ولكن عندما يصبح عدد المتغيِّرات أكبر من ذلك فعندئذ يتحدث المرء عن التّأثير المتبادل بين هذه المتغيِّرات على أوجه مختلفة، فإمَّا أن يُبْحث في التّأثير المتبادل بين كل هذه المتغيِّرات، فيكون المرء أمام ما يُعرف باسم **الارتباط المُتعدِّد المتغيِّرات** (أو الارتباط الكلي)، أو أن يُبْحث في التّأثير المتبادل بين جزء من هذه المتغيِّرات مع تثبيت بقيّة المتغيِّرات، فيكون المرء أمام ما يُعرف باسم **الارتباط الجزئي**. أمَّا بالنسبة إلى التغاير فيصبح لها شكل مصفوفة مربَّعة من مرتبة تساوي عدد المتغيِّرات وتُدعى **مصفوفة التغاير**.

### (٣,٤,١) الارتباط المتعدِّد Multiple Correlation

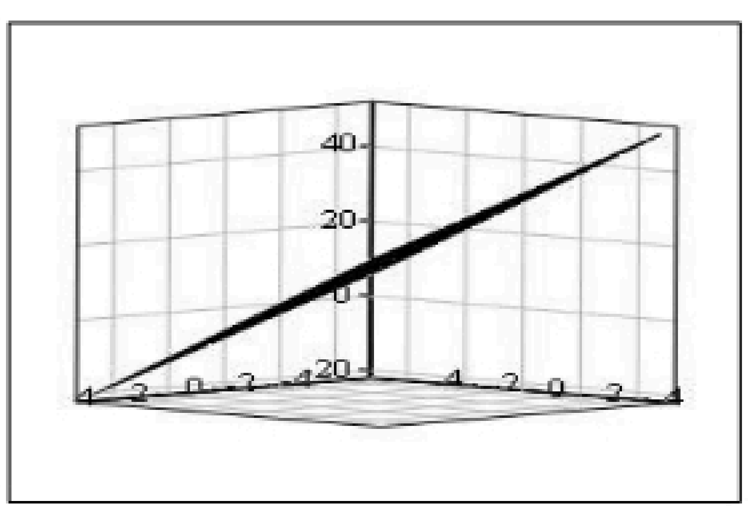
لقد ذكرنا سابقاً مقياساً للارتباط بين ظاهرتين فقط، إلاّ أنَّ كثيراً ما تكون الظاهرة التي هي قيد الدرس متأثّرة بظواهر كثيرة أخرى، فعلى سبيل المثال غالباً ما يكون سعر سلعة متأثّراً بأكثر من عامل، مثل أسعار سلع أخرى، وتوفّر هذه السلعة من عدمه، وإمكانية حصول المستهلك عليها، ومن ثمَّ فإنَّنا بحاجة لدراسة الارتباط بين أكثر من متغيِّرين كما هو مُلاحظ، وعلى العموم فلو افترضنا أنَّ لدينا $n$ متغيِّراً $X_1$، $X_2$، ... و $X_n$ تصف $n$ ظاهرة، وأنَّ العلاقة التي تربط بين المتغيِّر $X_k$ وبقيّة المتغيِّرات هي:

$$X_k = f\left( X_1, X_2, \ldots, X_{k-1}, X_{k+1}, \ldots, X_n \right) \qquad [3,34]$$

علماً أنَّ هذه العلاقة قد تكون علاقة خطيّة من الشكل:

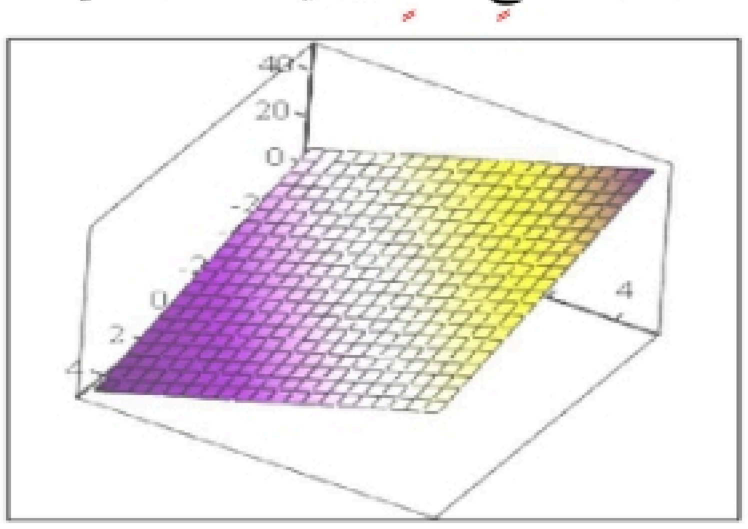
$$X_k = a_o + a_1 X_1 + a_2 X_2 + \ldots + a_{k-1} X_{k-1} + a_{k+1} X_{k+1}, + \ldots + a_n X_n$$

فعلى سبيل المثال لو كان لدينا ثلاثة متغيِّرات فقط، وكانت العلاقة الخطيّة بينها من الشكل $X_3 = 15 - 3X_1 + 4X_2$، فعندئذ سيكون تمثيل قيم هذه المتغيِّرات في الفضاء ثلاثي الأبعاد هي نقاط تقع على أو بالقرب من سطح مستوٍ كما في الشكل الآتي:




M

الشكل (٣,٣١.أ)



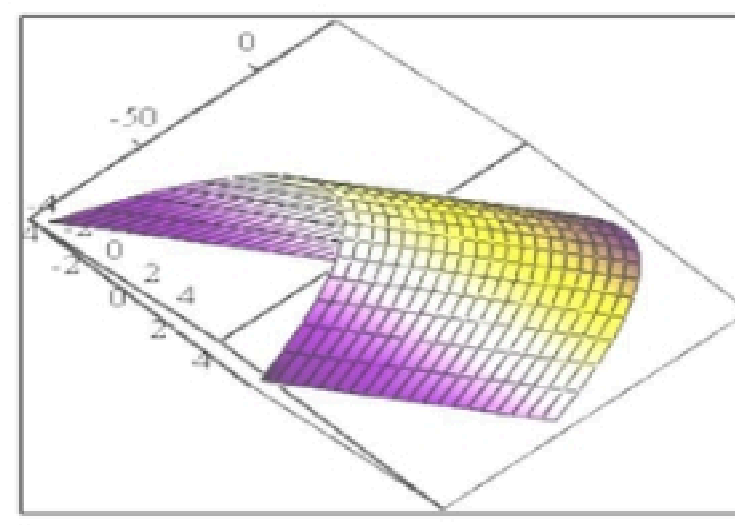

M

الشكل (٣,٣١.ب)

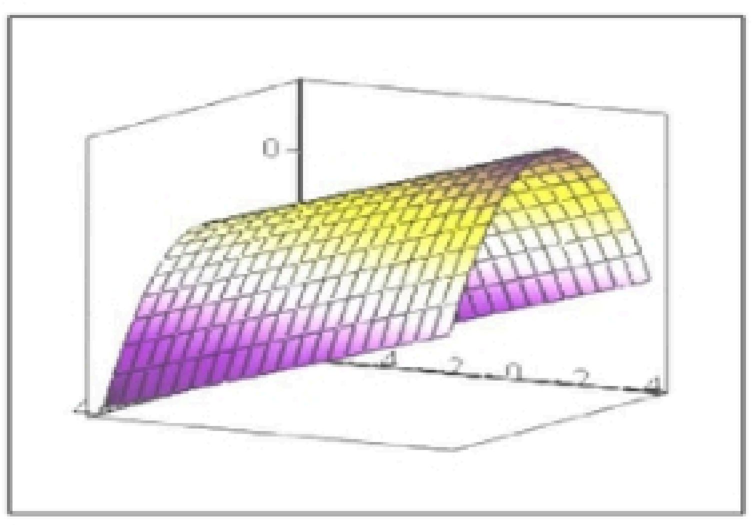
ولكن قد تكون العلاقة [3,34] غير خطيّة، فعلى سبيل المثال لو كان لدينا ثلاثة متغيِّرات فقط، وكانت العلاقة بينها من الشكل:

$$X_3 = 5 - 3X_1^2 + 5X_2$$

فعندئذ سيكون تمثيل قيم هذه المتغيِّرات في الفضاء ثلاثي الأبعاد هي نقاط تقع على أو بالقرب من سطحٍ مقوَّس كما في الشكل الآتي:




الشكل (٣٢,٣. ب)




الشكل (٣٢,٣.أ)

وهنا يلاحظ إمكانية قياس الارتباط بين $X_k$ وكلّ المتغيِّرات الأخرى مُجتمعة (أي تلك المتغيِّرات التي تتغيَّر كلها في وقت واحد) وذلك باستخدام أحد مقاييس الارتباط المتعدِّد.

الآن لتوضيح ما سبق سنقوم بدراسة الحالة البسيطة المتمثِّلة بالارتباط المتعدِّد الخطي وذلك عندما يكون لدينا ثلاثة متغيِّرات $X$، $Y$ و$Z$ فقط، وأنَّ المشاهدات الناتجة عن هذه المتغيِّرات هي $(x_1, y_1, z_1)$، $(x_2, y_2, z_2)$، ... و $(x_n, y_n, z_n)$، وأنَّه لدينا:

$$Z = a + bX + cY \qquad [3,35]$$

علماً أنَّ $a$، $b$ و$c$ في العلاقة [3–35] هي ثوابت يُطلب تقديرها حتى يمكننا النظر إلى السطح الممثَّل بالعلاقة [3–35] كأفضل توفيق للمشاهدات $(x_i, y_i, z_i)$ من أجل كل $i \in \mathbb{N}_n$.

إنَّ المبادئ العامّة للارتباط المتعدِّد هي نفسها كما هو الحال لدى الارتباط البسيط، إذ يمكننا استخدام طريقة المرّبعات الصغرى لكي نوفِّق سطحاً يصف العلاقة المتوسطة بين المتغيِّرات التي هي قيد الدراسة، وبحيث يكون مجموع انحرافات المشاهدات عن هذا السطح يساوي الصفر، وكذلك مجموع مربعات تلك الانحرافات عن ذلك السطح أقل ما يمكن. بعبارة أخرى، إنَّ المبادئ التي تُستخدم فيها طريقة المرّبعات الصغرى لدى توفيق خط يصف العلاقة المتوسطة بين متغيِّرين تنطبق على حالتنا هذه أيضاً، ولكن الفارق هو أنَّ لدينا في هذه الحالة ثلاثة محاور ونحاول توفيق سطح بين تلك المشاهدات بحيث يُحقِّق الشروط السابقة نفسها.

والآن لتقدير الثوابت $a$، $b$ و$c$ وفقاً لطريقة المرّبعات الصغرى نحصل على العلاقات الآتية:

$$\left.\begin{aligned}
\sum_{i=1}^{n} z_i &= a \cdot n + b\sum_{i=1}^{n} x_i + c\sum_{i=1}^{n} y_i \\
\sum_{i=1}^{n} x_i z_i &= a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 + c\sum_{i=1}^{n} x_i y_i \\
\sum_{i=1}^{n} y_i z_i &= a\sum_{i=1}^{n} y_i + b\sum_{i=1}^{n} x_i y_i + c\sum_{i=1}^{n} y_i^2
\end{aligned}\right\} \qquad [3,36]$$

وبالحل المشترك لجملة المعادلات [3–36] نحصل على قيم الثوابت $a$، $b$ و$c$، ومن ثمَّ التعويض في العلاقة السابقة [3–36] نحصل على العلاقة التي تقدِّر لنا قيمة $Z$ لدى تغيُّر قيمة $X$ و$Y$ معاً، وهذه العلاقة تُكتب بالشكل الآتي:

$$\hat{Z} = a + bX + cY$$

وبعد ذلك نقوم بحساب تباين القيم المشاهدة عن السطح الذي تمَّ توفيقه (المقدَّر للقيم المتوقعة ) كما يُفعل في الارتباط البسيط، وأمّا **الخطأ المعياري للتقدير** Standard Error الذي سنرمز له ب $S_{est\,\hat{Z}}$ ، فإنَّه يُحسب من خلال العلاقة الآتية:

$$S_{est\,\hat{Z}} = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n}\left(z_i - \hat{z}_i\right)^2} \qquad [3,37]$$

ومن ثمَّ تكون قيمة مُعامل الارتباط المتعدِّد بين $Z$ وكلٍّ من $X$ و$Y$ مُعطى من خلال العلاقة الآتية:

$$\mathbf{r}_{Z\,;\,X\,Y} = \sqrt{1-\frac{S^2_{est\,\hat{Z}}}{S^2_Z}} \qquad [3,38]$$

علماً أنَّ $S^2_{est\,\hat{Z}}$ هو تباين القيم المقدَّرة، وأمّا $S^2_Z$ فهو تباين القيم الفعليَّة للمتغيِّر $Z$.

### (٣, ٤, ١, ١) ملاحظات

١- إنَّ مُعامل الارتباط المتعدِّد يقيس العلاقة بين المتغيِّر التابع $Z$ والتغيُّر في المتغيِّرات المستقلَّة $X$ و$Y$ مُجتمعة.

٢- إنَّ القيمة المطلقة لمعامل الارتباط المتعدِّد لا يمكن لها أن تزيد عن الواحد، وكلّما اقتربت قيمته من الواحد؛ دلَّ ذلك على قوة الارتباط الخطي بين المتغيِّرات التي هي قيد الدرس، فإذا أصبح الارتباط خطيًّا تامًّا فعندئذ تصبح قيمة هذا المُعامل تساوي الواحد تماماً، وفي هذه الحالة ستصبح جميع النقاط الممثِّلة للمشاهدات واقعة على السطح الذي تمَّ توفيقه للبيانات. أمّا إذا كان مُعامل الارتباط يساوي الصفر، فعندئذ يجب ألاّ يُفسَّر ذلك بعدم وجود علاقة ارتباط بين المتغيِّرات التي هي قيد الدراسة، وإنَّما يعني ذلك عدم وجود علاقة خطيَّة بينها على الأقل (إذ إنَّه من الممكن أن تكون هناك نقاط شاذّة بسلوكها أيضاً أدّت إلى هذه النتيجة، كما شاهدنا ذلك سابقاً).

٣- في الحالة التي يكون فيها $n$ ظاهرة ممثَّلة بـ $n$ متغيِّر $X_1$، $X_2$، ... و $X_n$، وأنَّ العلاقة التي تربط بين متغيِّر $X_k$ وبقيّة المتغيِّرات هي علاقة خطيّة، فعندئذ لتقدير أمثال الانحدار $a_o$، $a_1$، $a_2$، ...و $a_n$، يجب علينا تعيين الحل المشترك لجملة المعادلات الآتية:

$$\sum_{i=1}^{n}x_{k,i} = a_o\cdot n + a_1\sum_{i=1}^{n}x_{1,i} + \dots + a_{k-1}\sum_{i=1}^{n}x_{k-1,i} + a_{k+1}\sum_{i=1}^{n}x_{k+1,i} + \dots + a_n\sum_{i=1}^{n}x_{n,i}$$

$$\sum_{i=1}^{n}x_{1,i}\cdot x_{k,i} = a_o\sum_{i=1}^{n}x_{1,i} + a_1\sum_{i=1}^{n}x^2_{1,i} + a_2\sum_{i=1}^{n}x_{1,i}\cdot x_{2,i} + \dots + a_{k-1}\sum_{i=1}^{n}x_{1,i}\cdot x_{k-1,i} + a_{k+1}\sum_{i=1}^{n}x_{1,i}\cdot x_{k+1,i} + \dots + a_n\sum_{i=1}^{n}x_{1,i}\cdot x_{n,i}$$

$$\sum_{i=1}^{n}x_{2,i}\cdot x_{k,i} = a_o\sum_{i=1}^{n}x_{2,i} + a_1\sum_{i=1}^{n}x_{2,i}\cdot x_{1,i} + a_2\sum_{i=1}^{n}x^2_{2,i} + \dots + a_{k-1}\sum_{i=1}^{n}x_{2,i}\cdot x_{k-1,i} + a_{k+1}\sum_{i=1}^{n}x_{2,i}\cdot x_{k+1,i} + \dots + a_n\sum_{i=1}^{n}x_{2,i}\cdot x_{n,i}$$

.................................................................

$$\sum_{i=1}^{n}x_{n,i}\cdot x_{k,i} = a_o\sum_{i=1}^{n}x_{n,i} + a_1\sum_{i=1}^{n}x_{n,i}\cdot x_{1,i} + a_2\sum_{i=1}^{n}x_{n,i}\cdot x_{2,i} + \dots + a_{k-1}\sum_{i=1}^{n}x_{n,i}\cdot x_{k-1,i} + a_{k+1}\sum_{i=1}^{n}x_{n,i}\cdot x_{k+1,i} + \dots + a_n\sum_{i=1}^{n}x^2_{n,i}$$

### (٣, ٤, ١, ٢) مثال

لنفترض أنَّه لدينا البيانات الآتية التي فيها $X$ يُمثِّل عدد الساعات الأسبوعية التي يُمضيها 15 طالباً في دراسة مقرّري الرياضيات والإحصاء في المنزل (بعد انتهاء الدوام الجامعي)، وأمّا $Y$ و$Z$ فيمثِّلان درجة الاختبار النهائي (علماً أنَّ الدرجة الكاملة **60**) في مقرّري الرياضيات والإحصاء على الترتيب.

الجدول (٢٣, ٣. أ)

| $i$ | $x_i$ | $y_i$ | $z_i$ | $i$ | $x_i$ | $y_i$ | $z_i$ |
|---|---|---|---|---|---|---|---|
| 1 | 0 | 21 | 15 | 9 | 2 | 25 | 30 |
| 2 | 0 | 18 | 15 | 10 | 2 | 38 | 45 |
| 3 | 0 | 22 | 21 | 11 | 3 | 44 | 50 |
| 4 | 1 | 24 | 28 | 12 | 3 | 51 | 60 |
| 5 | 1 | 25 | 30 | 13 | 4 | 39 | 45 |
| 6 | 1 | 25 | 35 | 14 | 4 | 54 | 60 |
| 7 | 1 | 26 | 40 | 15 | 5 | 55 | 50 |
| 8 | 2 | 34 | 35 | | | | |

ولنفترض أنّنا نرغب بتعيين العلاقة الخطّية التي تربط بين نتائج اختبار مقرَّر الإحصاء وكلٍّ من نتائج مقرَّر الرياضيات والمدّة الزمنية التي أمضاها لدراسة هذين المقرَّرين. في هذه الحالة ستكون العلاقة بين هذه المتغيِّرات من الشكل:

$$Z = a + bX + cY \qquad (1)$$

ولكن لتقدير الأمثال $a$، $b$ و $c$ علينا حساب المجاميع المُعطاة بالعلاقات [3–36] من أجل $n = 15$، حيث لدينا:

الجدول (٢٣, ٣. ب)

| $i$ | $x_i$ | $y_i$ | $z_i$ | $x_i \cdot y_i$ | $x_i \cdot z_i$ | $y_i \cdot z_i$ | $x_i^2$ | $y_i^2$ | $z_i^2$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 21 | 15 | 0 | 0 | 315 | 0 | 441 | 225 |
| 2 | 0 | 18 | 15 | 0 | 0 | 270 | 0 | 324 | 225 |
| 3 | 0 | 22 | 21 | 0 | 0 | 462 | 0 | 484 | 441 |
| 4 | 1 | 24 | 28 | 24 | 28 | 672 | 1 | 576 | 784 |
| 5 | 1 | 25 | 30 | 25 | 30 | 750 | 1 | 625 | 900 |
| 6 | 1 | 25 | 35 | 25 | 35 | 875 | 1 | 625 | 1225 |
| 7 | 1 | 26 | 40 | 26 | 40 | 1040 | 1 | 676 | 1600 |
| 8 | 2 | 34 | 35 | 38 | 70 | 1190 | 4 | 1156 | 1225 |
| 9 | 2 | 25 | 30 | 50 | 60 | 750 | 4 | 625 | 900 |
| 10 | 2 | 38 | 45 | 76 | 90 | 1710 | 4 | 1444 | 2025 |
| 11 | 3 | 44 | 50 | 132 | 150 | 2200 | 9 | 1936 | 2500 |
| 12 | 3 | 51 | 60 | 153 | 180 | 3060 | 9 | 2601 | 3600 |
| 13 | 4 | 39 | 45 | 156 | 180 | 1755 | 16 | 1521 | 2025 |
| 14 | 4 | 54 | 60 | 216 | 240 | 3240 | 16 | 2916 | 3600 |
| 15 | 5 | 55 | 50 | 275 | 250 | 2750 | 25 | 3025 | 2500 |
| sum | 29 | 501 | 559 | 1226 | 1353 | 21039 | 91 | 18975 | 23775 |

وبعد التعويض عن قيم المجاميع في المعادلات [3–36] كلّ بما يساويه من الجدول السابق (٢٣, ٣. ب) نحصل على جملة المعادلات الخطّية الآتية:

$$\left.\begin{aligned} 559 &= 15a + 29b + 501c \\ 21039 &= 29a + 91b + 1226c \\ 21039 &= 501a + 1226b + 18975c \end{aligned}\right\} \qquad (2)$$

وبحل جملة المعادلات (2) نجد أنَّ مقدَّرات الأمثال $a$، $b$ و $c$ هي $a = 2.08$، $b = 0.06$ و $c = 1.05$ على الترتيب، ومن ثَمَّ ستكون لعلاقة الارتباط (1) العرض الآتي:

$$\hat{Z} = 2.08 + 0.06X + 1.05Y$$

ومن أجل الحصول على قيمة مُعامل الارتباط المتعدِّد سنقوم أولاً بحساب الخطأ المعّياري للتقدير حيث لدينا:

$$\hat{z}_i = 2.08 + 0.06\ x_i + 1.05\ y_i$$

ومنه يكون:

الجدول (٢٣ ,٣. ج)

| $i$ | $z_i$ | $\hat{z}_i$ | $(z_i - \hat{z}_i)^2$ |
|---|---|---|---|
| 1 | 15 | 24.13 | 83.3569 |
| 2 | 15 | 20.98 | 35.7604 |
| 3 | 21 | 25.18 | 17.4724 |
| 4 | 28 | 27.34 | 0.4356 |
| 5 | 30 | 28.39 | 2.5921 |
| 6 | 35 | 28.39 | 43.6921 |
| 7 | 40 | 29.44 | 111.5136 |
| 8 | 35 | 37.9 | 8.41 |
| 9 | 30 | 28.45 | 2.4025 |
| 10 | 45 | 42.1 | 8.41 |
| 11 | 50 | 48.46 | 2.3716 |
| 12 | 60 | 55.81 | 17.5561 |
| 13 | 45 | 43.27 | 2.9929 |
| 14 | 60 | 59.02 | 0.9604 |
| 15 | 50 | 60.13 | 102.6169 |
| sum | 559 | 558.99 | 440.5435 |

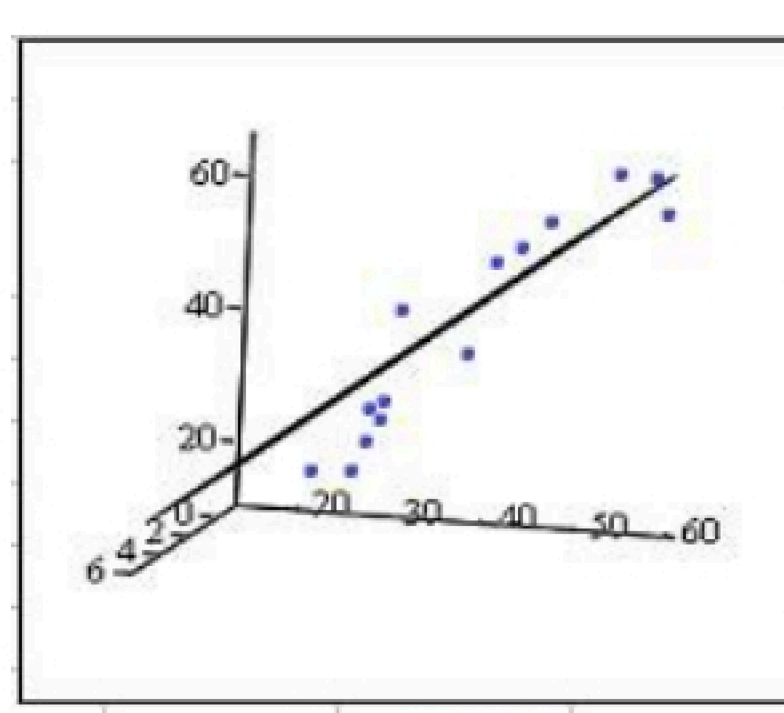
وبحسب العلاقة [37–3] تكون قيمة الخطأ المعياري للتقدير تساوي:

$$S_{est\,\hat{Z}} = \sqrt{\frac{1}{n-2}\sum_{i=1}^{n}\left(z_i - \hat{z}_i\right)^2} = \sqrt{\frac{440.5435}{13}} = 5.821$$

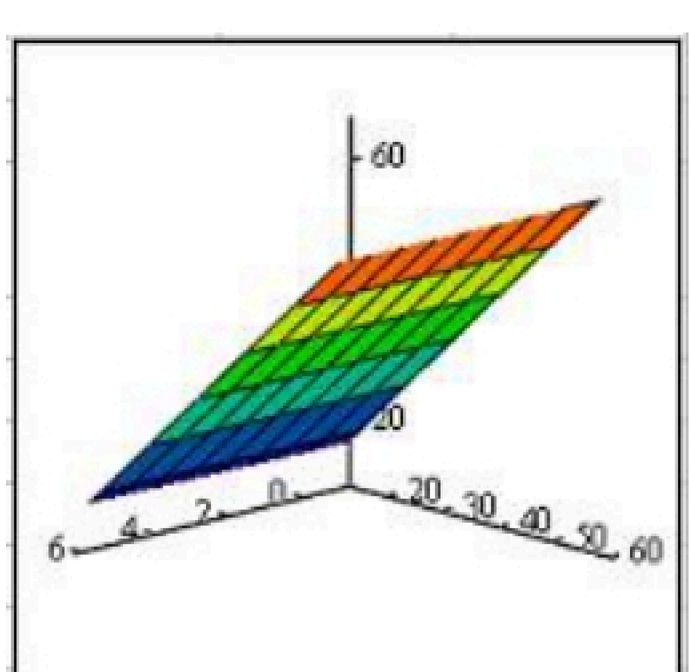
ومن ثَمّ تكون قيمة مُعامل الارتباط المتعدِّد هي:

$$\mathbf{r}_{Z\,;\,X\,Y} = \sqrt{1 - \frac{S^2_{est\,\hat{Z}}}{S^2_Z}} = \sqrt{1 - \frac{33.888}{210.2095}} = 0.839$$

وهذا يعني وجود ارتباط قوي جداً بين نتائج اختبار مقرَّر الإحصاء من جانب وتمكّن الطالب في الرياضيات مع المدّة التي يُمضيها في دراسة هذين المقرّرين. انظر الشكل الآتي الذي يبيِّن المستوي الموفّق للمشاهدات والتوضّع النسبي للنقاط الممثّلة للمشاهدات مع هذا المستوي.




الشكل (٣٣ ,٣. أ)




الشكل (٣٣ ,٣. ب)

**(٣ , ٤ , ١ , ٣) ملاحظة**

نلاحظ مما سبق أنَّ هذه الطريقة تعطي طريقة عامّة لحساب قيمة مُعامل الارتباط المتعدِّد، وحتى في حال أصبح عدد المتغيِّرات المستقلّة أكثر من ثلاثة، فإنَّه يُبحث في تقدير الثوابت التي تحدِّد لنا أفضل شكل ممثِّل للعلاقة المتوسطة بين المشاهدات، ومن ثمَّ حساب الخطأ المعياري مع أخذ المتغيِّر الجديد في الحسبان لدى حساب الخطأ المعياري المستخدم، وبعد ذلك حساب مُعامل الارتباط باستخدام العلاقة السّابقة نفسها. لكن وجد أنَّ مُعامل الارتباط المتعدِّد قابل للحساب من خلال معرفة مُعاملات الارتباط البسيط أيضاً وهذا ما ستوضِّحه لنا الفقرة الآتية.

**(٢ , ٤ , ٣) العلاقة بين الارتباط المتعدِّد والارتباط البسيط**

لقد رأينا أنَّ العلاقة التي تُعطي الارتباط المتعدِّد بين ثلاثة متغيِّرات $X$، $Y$ و$Z$ هي العلاقة [3,38]، علماً أنَّ المتغيِّر $Z$ هو المتغيِّر التابع وأمّا $X$ و$Y$ فهما المتغيِّران المستقلّان، ولكن في الواقع يمكننا الحصول على قيمة مُعامل الارتباط المتعدِّد إذا علمنا مُعاملات الارتباط البسيطة بين كل زوج من المتغيِّرات $X$، $Y$ و$Z$، فعلى سبيل المثال:

١- إذا كان لدينا ثلاثة متغيِّرات $X_1$، $X_2$ و $X_3$ فعندئذ من أجل أي تبديل $(i_1\, i_2\, i_3)$ للأعداد 1، 2 و3 تكون العلاقة الآتية مُحقَّقة:

$$\mathbf{r}_{i_1;i_2,i_3} = \sqrt{\frac{\mathbf{r}^2_{i_1,i_2} + \mathbf{r}^2_{i_1,i_3} - 2\,\mathbf{r}^2_{i_1,i_2}\ \mathbf{r}^2_{i_1,i_3}\ \mathbf{r}^2_{i_2,i_3}}{1-\mathbf{r}^2_{i_2,i_3}}} \qquad [3,39]$$

وهكذا نلاحظ أنَّ قيمة مُعامل الارتباط المتعدِّد أكبر أو تساوي قيمة مُعاملات الارتباط البسيطة. كما يمكن حساب $\mathbf{r}_{i_1;i_2,i_3}$ من خلال إحدى العلاقتين الآتيتين أيضاً:

$$\mathbf{r}_{i_1;i_2,i_3} = \sqrt{\mathbf{r}^2_{i_1,i_2} + \mathbf{r}^2_{i_1,i_2;i_3} - \left(1-\mathbf{r}^2_{i_1,i_2}\right)} \qquad [3,40]$$

أو

$$\mathbf{r}_{i_1;i_2,i_3} = \sqrt{1-\left(1-\mathbf{r}^2_{i_1,i_2}\right)\cdot\left(1-\mathbf{r}^2_{i_1,i_3;i_2}\right)} \qquad [3,41]$$

٢- إذا كان لدينا أربعة متغيِّرات $X_1$، $X_2$، $X_3$ و $X_4$ فعندئذ من أجل أي تبديل $(i_1\, i_2\, i_3\, i_4)$ للأعداد 1، 2، 3 و4 تكون العلاقة الآتية مُحقَّقة:

$$\mathbf{r}_{i_1;i_2,i_3,i_4} = \sqrt{1-\left(1-\mathbf{r}^2_{i_1,i_2}\right)\cdot\left(1-\mathbf{r}^2_{i_1,i_3;i_2}\right)\cdot\left(1-\mathbf{r}^2_{i_1,i_4;i_2,i_3}\right)} \qquad [3,42]$$

٣- إذا كان لدينا $k$ متغيِّر $X_1$، $X_2$، ... و $X_k$، فعندئذ من أجل أي تبديل $(i_1\, i_2\, i_3 \ldots i_k)$ للأعداد 1، 2، 3، ... و$k$ ستكون العلاقة الآتية مُحقَّقة:

$$\mathbf{r}_{i_1;i_2,i_3,\ldots,i_k} = \sqrt{1-\left(1-\mathbf{r}^2_{i_1,i_2}\right)\cdot\left(1-\mathbf{r}^2_{i_1,i_3;i_2}\right)\cdot\ldots\cdot\left(1-\mathbf{r}^2_{i_1,i_k;i_2,i_3,\ldots,i_{k-1}}\right)} \qquad [3,43]$$

**(١ , ٢ , ٤ , ٣) مثال**

لنأخذ مجموعة البيانات الآتية التي تمثِّل تطور الطول $Y$ والوزن $Z$ مع العمر $X$ للأطفال منذ الولادة وحتى الشهر 162 من عمره (هذه البيانات أخذت عن كتاب طب الأطفال – نشر جامعة دمشق 1988–1989)، فكانت النتائج كما في الجدول الآتي:

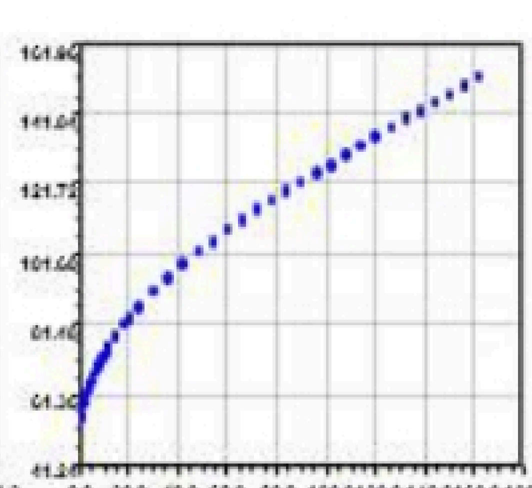
الجدول (٣,٢٤)

| العمر $X$ بالشهر | الطول $Y$ سم | الوزن $Z$ كغ |
|---|---|---|
| 0 | 54.3 | 3.36 |
| 1 | 55.6 | 4.72 |
| 2 | 58.7 | 5.44 |
| 3 | 61.2 | 6.17 |
| 4 | 63.5 | 6.80 |
| 5 | 65.3 | 7.17 |
| 6 | 67.1 | 7.85 |
| 7 | 68.8 | 8.16 |
| 8 | 70.1 | 8.48 |
| 9 | 71.4 | 8.80 |
| 10 | 72.6 | 9.07 |
| 11 | 73.9 | 9.39 |
| 12 | 74.9 | 9.71 |

| العمر $X$ بالشهر | الطول $Y$ سم | الوزن $Z$ كغ |
|---|---|---|
| 15 | 78 | 10.29 |
| 18 | 81 | 11.20 |
| 21 | 83.6 | 11.80 |
| 24 | 86.1 | 12.40 |
| 30 | 90.7 | 13.20 |
| 36 | 94.7 | 14.50 |
| 42 | 98.6 | 15.40 |
| 48 | 102.1 | 16.10 |
| 54 | 105.4 | 17.10 |
| 60 | 108.5 | 17.80 |
| 66 | 111.5 | 19.00 |
| 72 | 114.3 | 19.90 |
| 78 | 117.1 | 20.80 |

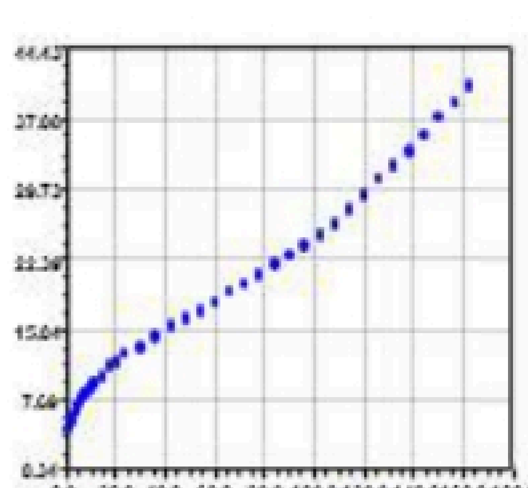
| العمر $X$ بالشهر | الطول $Y$ سم | الوزن $Z$ كغ |
|---|---|---|
| 84 | 119.9 | 21.80 |
| 90 | 122.4 | 22.90 |
| 96 | 125.0 | 23.90 |
| 102 | 127.5 | 25.10 |
| 108 | 130.0 | 26.30 |
| 114 | 132.6 | 27.70 |
| 120 | 135.1 | 29.20 |
| 126 | 137.7 | 30.70 |
| 132 | 140.2 | 32.30 |
| 138 | 142.7 | 33.90 |
| 144 | 145.0 | 35.50 |
| 150 | 147.3 | 37.20 |
| 156 | 149.6 | 38.90 |
| 162 | 151.9 | 40.70 |

فلو أخذنا $X$ متغيِّر يرصد الشهر الذي أخذ فيه القياس مع $X = 0$ هي لحظة الولادة، $Y$ متغيِّر يرصد طول الطفل و$Z$ متغيِّر يرصد وزن الطفل، فعندئذ سنحصل على العروض الآتية التي تُظهر لنا توزع النقاط الممثِّلة لقيم هذه المتغيِّرات مأخوذةً مثنى مثنى:



الشكل (٣,٣٤. أ)
علاقة الوزن بالطول
$Z = f(Y)$

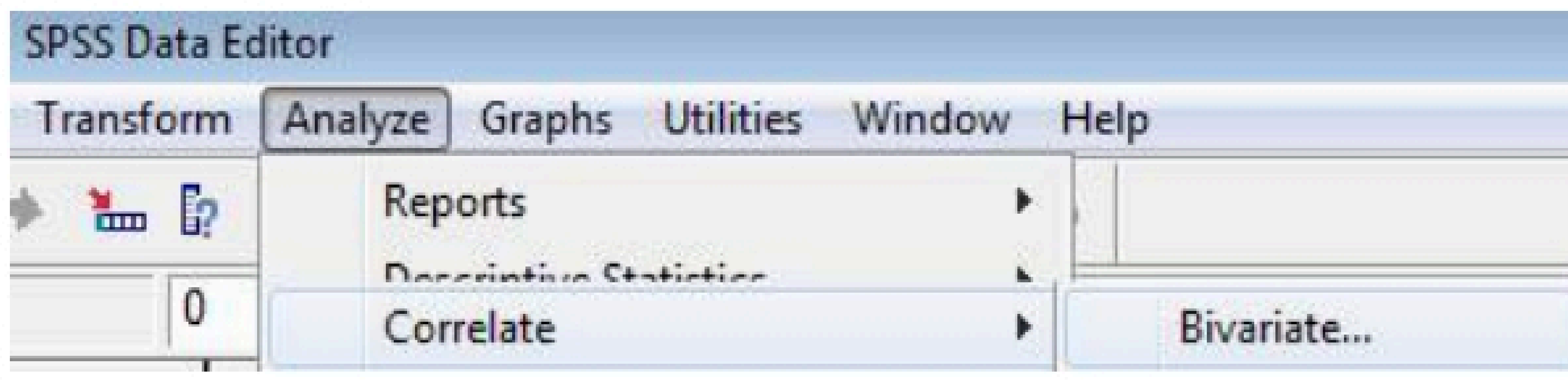
الشكل (٣,٣٤. ب)
تطور الوزن مع العمر
$Z = f(X)$

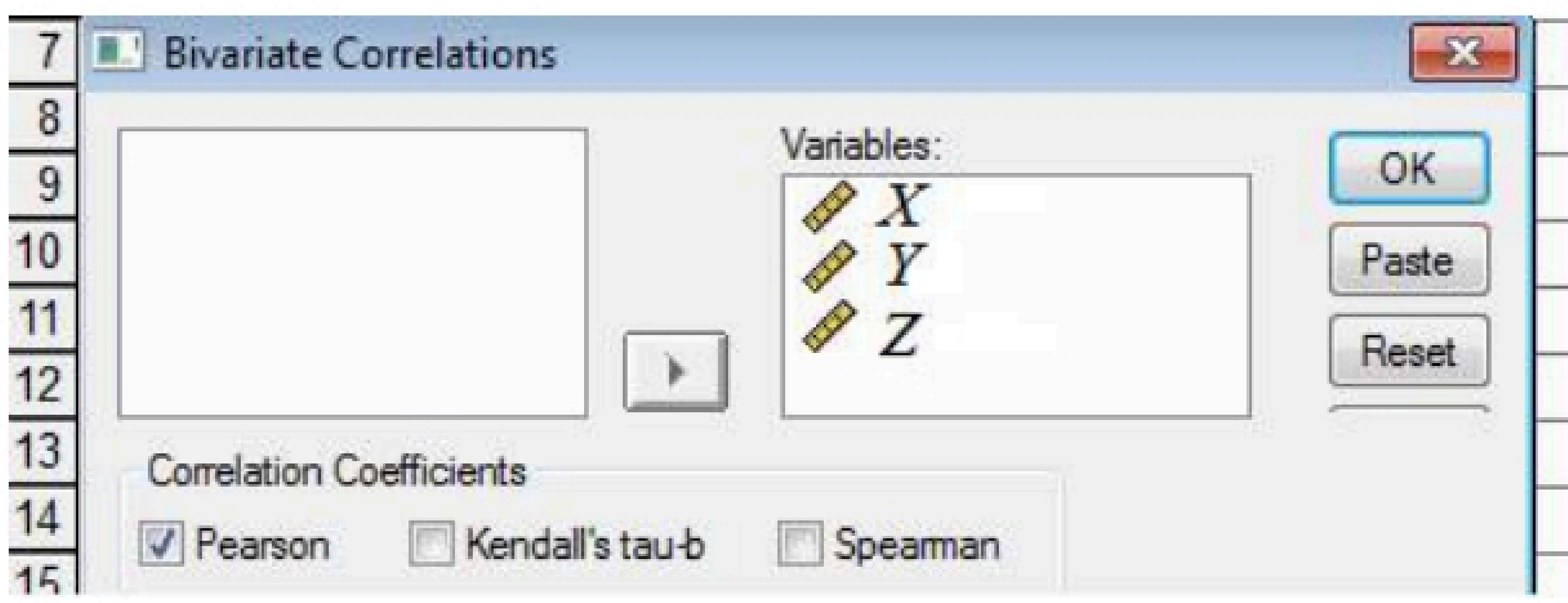


الشكل (٣,٣٤. ج)
تطور الطول مع العمر
$Y = f(X)$

و هكذا نلاحظ أنَّه ثمَّة علاقة تربط بين كل متغيِّرين على حدة، وأنَّ هذه العلاقات تتمايز فيما بينها.

الآن لحساب مُعاملات الارتباط المتعدِّد (الكلي) لهذه المتغيِّرات الثلاثة سنقوم أولاً بحساب مُعاملات الارتباط الخطي البسيط بين هذه المتغيِّرات، فنجد $\mathbf{r}_{XY} = 0.982$ ، $\mathbf{r}_{XZ} = 0.994$ و $\mathbf{r}_{YZ} = 0.981$ . بالطبع كان من الممكن استخدام برنامج **SPSS** للحصول على قيم الارتباط الخطي البسيط بين المتغيِّرات من خلال الخطوات الآتية:




فنحصل على التقرير الآتي:

**Correlations**

| | | X | Y | Z |
|---|---|---|---|---|
| X | Pearson Correlation | 1 | .982** | .994** |
| | Sig. (2-tailed) | | .000 | .000 |
| | N | 40 | 40 | 40 |
| Y | Pearson Correlation | .982** | 1 | .981** |
| | Sig. (2-tailed) | .000 | | .000 |
| | N | 40 | 40 | 40 |
| Z | Pearson Correlation | .994** | .981** | 1 |
| | Sig. (2-tailed) | .000 | .000 | |
| | N | 40 | 40 | 40 |

ومن ثمَّ سيكون لمعاملات الارتباط المتعدِّد القيم الآتية:

$$\mathbf{r}_{X;Y,Z} = \sqrt{\frac{\mathbf{r}_{XY}^2 + \mathbf{r}_{XZ}^2 - 2\mathbf{r}_{XY}\,\mathbf{r}_{XZ}\,\mathbf{r}_{YZ}}{1-\mathbf{r}_{YZ}^2}} = \sqrt{\frac{0.03724}{0.03764}} = \sqrt{0.989} = 0.995$$

$$\mathbf{r}_{Y;X,Z} = \sqrt{\frac{\mathbf{r}_{XY}^2 + \mathbf{r}_{YZ}^2 - 2\mathbf{r}_{XY}\,\mathbf{r}_{XZ}\mathbf{r}_{YZ}}{1-\mathbf{r}_{XZ}^2}} = \sqrt{\frac{0.011561}{0.011964}} = \sqrt{0.966} = 0.983$$

$$\mathbf{r}_{Z;X,Y} = \sqrt{\frac{\mathbf{r}_{XZ}^2 + \mathbf{r}_{YZ}^2 - 2\mathbf{r}_{XY}\,\mathbf{r}_{XZ}\mathbf{r}_{YZ}}{1-\mathbf{r}_{XY}^2}} = \sqrt{\frac{0.035273}{0.035676}} = \sqrt{0.989} = 0.994$$

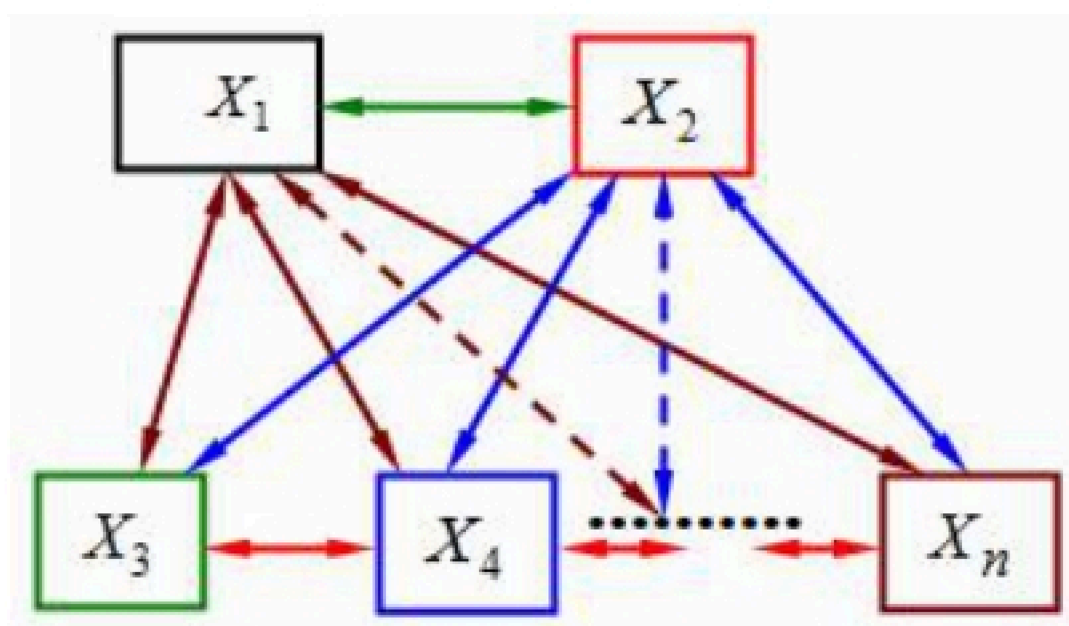
وبهذا يكون قد تمَّ حساب المُعاملات المذكورة آنفًا حيث نلاحظ القوة الكبيرة للارتباط في كل حالة من هذه الحالات الثلاث.

### (٣, ٤, ٣) الارتباط الجزئي Partial Correlation

في كثير من الأحيان قد يرتبط المتغيِّر التابع بالعديد من المتغيِّرات، ولكن يرغب المرء عادةً أن يدرس العلاقة بين المتغيِّر التابع وواحد فقط من هذه المتغيِّرات المستقلَّة مع عدم تجاهل المتغيِّرات الأخرى (وكأنه يبحث في إيجاد مُعامل الارتباط البسيط بين المتغيِّر التابع ومتغيِّر مستقل واحد قيد الدرس). في مثل هذه الحالة يستخدم المرء مُعامل الارتباط الجزئي الذي يقيس العلاقة بين المتغيِّر التابع وأحد المتغيِّرات المستقلَّة مُفترضاً ثبات تأثير بقيّة المتغيِّرات المستقلَّة الأخرى عند مستوى مُحدَّد، والفقرة الآتية تشرح لنا هذا المفهوم.

#### (٣, ٤, ٣, ١) تعريف (الارتباط الجزئي)

لتكن لدينا المتغيِّرات $X_1$، $X_2$، ... و $X_n$، فعندئذ يُعرَّف **الارتباط الجزئي** على أنَّه العلاقة المتبادلة الخطّيّة بين متغيِّرين من هذه المتغيِّرات لدى ارتباط كل المتغيِّرات الأخرى بتأثيراتها المتبادلة فيما بينها كما في الشكل التشابكي الآتي (٣, ٣٨).




الشكل (٣, ٣٨)

من أجل توضيح هذا المفهوم سنأخذ الحالة التي يكون اهتمامنا فيها مُنصبًّا على مدى الارتباط الفعلي بين $X_1$ و $X_2$ تحت التأثيرات الناتجة عن المتغيِّرات $X_1$، $X_2$، ... و $X_n$، وكذلك سنرمز للارتباط الجزئي بين $X_1$ و $X_2$ تحت التأثيرات الناتجة عن المتغيِّرات $X_3$ و $X_4$، ... و $X_n$ بالرمز $X_1\ X_2\ ;\ X_3\ X_4 \dots X_n$، وأمّا قيمة **مُعامل الارتباط الجزئي** له Partial Correlation Coefficient فسوف نرمز له بـ $\mathbf{r}_{1,2\,;\,3,4,\dots,n}$، علماً أنَّ العدد $n-2$ الذي يُمثِّل عدد المتغيِّرات $X_3$ و $X_4$، ... و $X_n$ يُدعى **رتبة الارتباط الجزئي** Order of the Partial Correlation، وأخيراً يُطلق على المتغيِّرات $X_3$ ، $X_4$، ... و $X_n$ اسم **المتغيِّرات المشوِّشة** Confusing Variables.

في الواقع إنَّ مُعامل الارتباط الجزئي $\mathbf{r}_{1,2\,;\,3,4,\dots,n}$ هو مقياس للارتباط بين الانحرافات الآتية (التي سنرمز لها بـ $X_{1;\,3,4,\dots,n}$):

$$\left(X_1-\overline{X_1}\right)-\alpha_{13}\left(X_3-\overline{X_3}\right)-\alpha_{14}\left(X_4-\overline{X_4}\right)-\dots-\alpha_{1n}\left(X_n-\overline{X_n}\right)$$

لقيم المتغيِّر $X_1$ عن دالَّة انحداره للمتغيِّرات $X_3$ و $X_4$، ... و $X_n$، وكذلك مقياس للارتباط بين الانحرافات الآتية (التي سنرمز لها بـ $X_{2;\,3,4,\dots,n}$):

$$\left(X_2-\overline{X_2}\right)-\alpha_{23}\left(X_3-\overline{X_3}\right)-\alpha_{24}\left(X_4-\overline{X_4}\right)-\dots-\alpha_{2n}\left(X_n-\overline{X_n}\right)$$

لقيم المتغيِّر $X_2$ عن دالَّة انحداره للمتغيِّرات $X_3$ و $X_4$، ... و $X_n$، وهكذا فإنَّ مُعاملات الارتباط الجزئي $\mathbf{r}_{1,2\,;\,3,4,\dots,n}$ شأنها شأن مُعامل الارتباط البسيط $\mathbf{r}_{1,2}$، أي إنَّه سيكون لدينا:

$$-1 \le \mathbf{r}_{1,2\,;\,3,4,\dots,n} \le +1 \qquad [3,44]$$

إنَّ المتغيِّرات $X_{1;3,4,\ldots,n}$ و $X_{2;3,4,\ldots,n}$ تُعبِّر عن تغيُّر $X_1$ و $X_2$ بالنسبة إلى المتغيِّرات المستبعدة $X_3$ و $X_4$، ... و $X_n$ التي يُفهم منها أنَّه يمكن أخذ $X_3$ و $X_4$، ... و $X_n$ كثوابت، وأنَّ القيم $\alpha_{1i}$ و $\alpha_{2i}$ من أجل كل $i = 3,4,\ldots,n$ هي التقديرات لمعاملات الارتباط الجزئي.

في الواقع إنَّ مُعامل الارتباط الجزئي $\mathbf{r}_{1,2;3,4,\ldots,n}$ يزودنا بالتحليل الإحصائي لسلوك الارتباط من أجل كل مشاهدتين (أو ملحوظتين) من مشاهدات متعدِّدة.

الآن، ومن أجل دراسة مُعامل الارتباط الجزئي سنميِّز بين الحالات الآتية:

١- إذا كان تأثير المتغيِّرات $X_3$ و $X_4$، ... و $X_n$ على $X_1$ و $X_2$ صغيراً لدرجة أنَّ الارتباط الجزئي $X_1\ X_2\ ;\ X_3\ X_4 \ldots X_n$ ليس له أهمية مقارنة بالارتباط البسيط بين $X_1$ و $X_2$، فعندئذ ستكون قيمة $\mathbf{r}_{1,2;3,4,\ldots,n}$ قريبةً جداً من قيمة $\mathbf{r}_{1,2}$، أي إنه ستكون لدينا:

$$\mathbf{r}_{1,2;3,4,\ldots,n} \approx \mathbf{r}_{1,2}$$

٢- إذا كانت العلاقة الارتباطية بين $X_1$ وتأثير المتغيِّرات $X_3$ و $X_4$، ... و $X_n$ قويةً بشكل كبير، وكذلك العلاقة الارتباطية بين $X_2$ وتأثير المتغيِّرات $X_3$ و $X_4$، ... و $X_n$ قويةً بشكلٍ كبير أيضاً، فعندئذ ستكون قيمة $\mathbf{r}_{1,2;3,4,\ldots,n}$ قريبةً من الصفر، أي إنَّه ستكون لدينا:

$$\mathbf{r}_{1,2;3,4,\ldots,n} \approx 0$$

٣- إنَّ استبعاد تأثير المتغيِّرات المشوِّشة يؤثِّر في العلاقة الارتباطية بين $X_1$ و $X_2$ والمعنيين بدراستها، وهذا الاستبعاد سيظهر الارتباط الجزئي $X_1\ X_2\ ;\ X_3\ X_4 \ldots X_n$ وكأنَّه الارتباط البسيط بين $X_1$ و $X_2$، وهذه الحالة تحدث بشكل خاص عندما يكون لأحد المتغيِّرات (أو لثلاثة منها أو لخمسة منها أو...) من الارتباطات المضاعفة مع المتغيِّرات المستبعدة له قيمة سالبةٌ.

على كل الأحوال، فإنَّ المدلول العمليّ لقيمة الارتباط الجزئي $X_1\ X_2\ ;\ X_3\ X_4 \ldots X_n$ تعني أنَّ القيمة $\mathbf{r}_{1,2;3,4,\ldots,n}$ ستكون أكبر من القيمة العمليّة للارتباط البسيط $\mathbf{r}_{1,2}$، وأمَّا لحساب قيمة $\mathbf{r}_{1,2;3,4,\ldots,n}$ فيمكن استخدام العلاقة التراجعية الآتية:

$$\mathbf{r}_{1,2;3,4,\ldots,n} = \frac{\mathbf{r}_{1,2;3,4,\ldots,n-1} - \mathbf{r}_{1,n;3,4,\ldots,n-1} \cdot \mathbf{r}_{2,n;3,4,\ldots,n-1}}{\sqrt{\left(1-\mathbf{r}^2_{1,n;3,4,\ldots,n-1}\right)\cdot\left(1-\mathbf{r}^2_{2,n;3,4,\ldots,n-1}\right)}} \qquad [3,45]$$

أو وفقاً للعلاقة التراجعية الآتية أيضاً:

$$\mathbf{r}_{1,2;3,4,\ldots,n} = \frac{\mathbf{r}_{1,2;4,5,\ldots,n} - \mathbf{r}_{1,3;4,5,\ldots,n} \cdot \mathbf{r}_{2,3;4,5,\ldots,n}}{\sqrt{\left(1-\mathbf{r}^2_{1,3;4,5,\ldots,n}\right)\cdot\left(1-\mathbf{r}^2_{2,3;4,5,\ldots,n}\right)}} \qquad [3,46]$$

### (٣,٤,٣,٢) ملاحظات

١- لاحظ أنَّ العلاقة السّابقة [3,45] تصبح غير معرَّفة عندما يكون لدينا:

$$\mathbf{r}_{1,n;3,4,\ldots,n-1} = \pm 1 \quad \text{or} \quad \mathbf{r}_{2,n;3,4,\ldots,n-1} = \pm 1$$

٢- عند استخدام العلاقة السابقة [3,45] يجب على المرء ألاَّ يرفع درجة الارتبـاط الجزئي كثيراً، بمعنى أنَّه عليه ألاَّ يستبعد الكثير من المتغيِّرات المشوِّشة لأنَّه يمكن إثبات أنَّ:

$$\mathbf{r}_{1,2\,;\,3,4,\ldots,\,n} < \mathbf{r}_{1,2\,;\,3,4,\ldots,\,n-1} \qquad [3,47]$$

وهكذا فإنَّه كلما كبرت قيمة $k$ أكثر فأكثر فإنَّ قيمة $\mathbf{r}_{1,2\,;\,3,4,\ldots,\,n}$ ستقترب من الصفر أكثر فأكثر، ومن ثمَّ من أجل $n$ كبيرة بقدرٍ كافٍ سيصبح لدينا:

$$\mathbf{r}_{1,2\,;\,3,4,\ldots,\,n} \approx 0$$

وهذا يعني بدوره أنَّه كلما استبعد المرء متغيِّرات مشوِّشة أكثر، فإنَّ ذلك سيؤدي إلى ضعف الحكم على الارتباط (أي إنه ليس لديه الكثير ما يقوله في هذه الصدد). إذ إنَّ ذلك الاستبعاد سيؤدي إلى تصغير حدود ما يُعرف باسم فترة الثقة المتعلِّقة بهذه المسألة، ولذلك لدى الدراسة والبحث في المسائل التطبيقية يجب مراعاة المتغيِّرات الأساسيّة جميعاً.

٣- في الحالة الخاصّة عندما يكون لدينا $k = 4$ نجد أنَّه يصبح للعلاقة التراجعية [3–45] العرض الآتي:

$$\mathbf{r}_{1,2\,;\,3,4} = \frac{\mathbf{r}_{1,2\,;\,3} - \mathbf{r}_{1,4\,;\,3} \cdot \mathbf{r}_{2,4\,;\,3}}{\sqrt{\left(1-\mathbf{r}^2_{1,4\,;\,3}\right)\cdot\left(1-\mathbf{r}^2_{2,4\,;\,3}\right)}} \qquad [3,48\text{-a}]$$

في حين يصبح للعـلاقة التراجعية [3–46] العرض الآتي:

$$\mathbf{r}_{1,2\,;\,3,4} = \frac{\mathbf{r}_{1,2\,;\,4} - \mathbf{r}_{1,3\,;\,4} \cdot \mathbf{r}_{2,3\,;\,4}}{\sqrt{\left(1-\mathbf{r}^2_{1,3\,;\,4}\right)\cdot\left(1-\mathbf{r}^2_{2,3\,;\,4}\right)}} \qquad [3,48\text{-b}]$$

وأمّا في الحالة الخاصّة عندما يكون لدينا $k = 3$ فإنَّنا نجد أن لكلٍّ من العلاقتين التراجعيتين [3,45] و [3,46]) العرض الآتي:

$$\mathbf{r}_{1,2\,;\,3} = \frac{\mathbf{r}_{1,2} - \mathbf{r}_{1,3}\ \mathbf{r}_{2,3}}{\sqrt{\left(1-\mathbf{r}^2_{1,3}\right)\cdot\left(1-\mathbf{r}^2_{2,3}\right)}} \qquad [3,49]$$

وهكذا نلاحظ أنَّ حساب العلاقة [3–48-a] و [3–48-b] سيؤول إلى استخدام العلاقة [3–49]، ومن ثمَّ استخدام مُعاملات الارتباط البسيط. لذلك عند حساب مُعاملات الارتباط الجزئي علينا أولاً تعيين جميع مُعاملات الارتباط البسيط ومن ثمَّ حساب مُعاملات الارتباط الجزئي $\mathbf{r}_{1,2\,;\,3,4,\ldots,\,n}$، وهذا بدوره سيتطلب جهداً كبيراً، ولذلك قُدِّمت طرائق أخرى لحساب مُعاملات الارتباط الجزئي $\mathbf{r}_{1,2\,;\,3,4,\ldots,\,n}$، ولكنَّنا لسنا بصدد تقديمها هنا ويمكن لمن يودّ الاطلاع عليها أن يعود إلى المراجع ذات الصلة بهذا الموضوع والمذكور بعض منها في فهرس المراجع في نهاية هذا الكتاب.

٤- بالعودة إلى الملاحظة السابقة (٣) من هذه الفقرة فإنَّه يمكن تبيان أنَّ:

$$\mathbf{r}_{1,2\,;\,3} = \sqrt{\alpha_{1,3} \cdot \alpha_{2,3}}$$

علماً أنَّ $\alpha_{1,3}$ هو مُعامل الانحدار لـ $X_1$ بالنسبة إلى $X_3$ لدى تثبيت $X_2$، وأمّا $\alpha_{2,3}$ فهو مُعامل الانحدار لـ $X_2$ بالنسبة إلى $X_3$ لدى تثبيت $X_1$.

الآن، وبشكلٍ مماثل لما سبق نجد أنَّ:

$$\mathbf{r}_{1,2;3,4} = \sqrt{\alpha_{1,2;3,4} \cdot \alpha_{2,1;3,4}}$$

علماً أنَّ $\alpha_{1,2;3,4}$ هو مُعامل انحـدار المتغيِّر $X_1$ بالنسبة إلى $X_2$ لدى تثبيت المتغيِّرين $X_3$ و $X_4$ ، وأمّا $\alpha_{2,1;3,4}$ فهو مُعامل انحدار المتغيِّر $X_2$ بالنسبة إلى $X_1$ لدى تثبيت المتغيِّرين $X_3$ و $X_4$ .

### (٣,٣,٤,٣) مثال

بالعودة إلى المثال (١-٢-٤-٣) حيث لدينا:

$$\mathbf{r}_{1,2} = \mathbf{r}_{XY} = 0.982 \quad \& \quad \mathbf{r}_{1,3} = \mathbf{r}_{XZ} = 0.994 \quad \& \quad \mathbf{r}_{2,3} = \mathbf{r}_{YZ} = 0.981$$

ومن ثمَّ سيكون:

$$\mathbf{r}_{1,2;3} = \frac{\mathbf{r}_{1,2} - \mathbf{r}_{1,3} \cdot \mathbf{r}_{2,3}}{\sqrt{\left(1-\mathbf{r}_{1,3}^2\right)\cdot\left(1-\mathbf{r}_{2,3}^2\right)}} = \frac{0.982 - 0.994 \times 0.981}{\sqrt{\left(1-0.994^2\right)\cdot\left(1-0.981^2\right)}} = 0.3245$$

وبالمثل نجد أنَّ:

$$\mathbf{r}_{1,3;2} = \frac{\mathbf{r}_{1,3} - \mathbf{r}_{1,2} \cdot \mathbf{r}_{3,2}}{\sqrt{\left(1-\mathbf{r}_{1,2}^2\right)\cdot\left(1-\mathbf{r}_{3,2}^2\right)}} = \frac{0.994 - 0.982 \times 0.981}{\sqrt{\left(1-0.982^2\right)\cdot\left(1-0.981^2\right)}} = 0.8366$$

وأخيراً يكون لدينا:

$$\mathbf{r}_{2,3;1} = \frac{\mathbf{r}_{2,3} - \mathbf{r}_{2,1} \cdot \mathbf{r}_{3,1}}{\sqrt{\left(1-\mathbf{r}_{2,1}^2\right)\cdot\left(1-\mathbf{r}_{3,1}^2\right)}} = \frac{0.981 - 0.982 \times 994}{\sqrt{\left(1-0.982^2\right)\cdot\left(1-0.994^2\right)}} = 0.2368$$

### (٣,٤,٤) العلاقة بين الارتباط المتعدِّد والارتباط الجزئي

يمكن للمرء استنتاج مُعاملات الارتباط المتعدِّد لدى معرفة مُعاملات الارتباط الجزئي، فعلى سبيل المثال إذا كان لدينا $n = 3$ فعندئذ من أجل $\mathbb{N}_3 \ni j, \ell, k$ مع $j \neq \ell \neq k$ فإنَّ مُعاملات الارتباط المتعدِّد بدلالة مُعاملات الارتباط الجزئي تُعطى بالعلاقة الآتية:

$$1-\mathbf{r}^2_{j;\ell,k} = \left(1-\mathbf{r}^2_{j,\ell}\right)\cdot\left(1-\mathbf{r}^2_{j,k;\ell}\right) \qquad [3,50]$$

وأمّا إذا كان لدينا $n = 4$ فعندئذ من أجل $\mathbb{N}_4 \ni j, \ell, k, m$ مع $j \neq \ell \neq k \neq m$ فإنَّ مُعاملات الارتباط المتعدِّد بدلالة مُعاملات الارتباط الجزئي تُعطى باستخدام العلاقة الآتية:

$$1-\mathbf{r}^2_{j;\ell,k,m} = \left(1-\mathbf{r}^2_{j,\ell}\right)\cdot\left(1-\mathbf{r}^2_{j,k;\ell}\right)\cdot\left(1-\mathbf{r}^2_{j,m;\ell,k}\right) \qquad [3,51]$$

كذلك يمكن حساب مُعاملات الارتباط الجزئي بدلالة مُعاملات الارتباط الكلي، فعلى سبيل المثال إذا كان لدينا $n = 3$ فعندئذ من أجل $\mathbb{N}_3 \ni j, \ell, k$ مع $j \neq \ell \neq k$ فإنَّ مُعاملات الارتباط الجزئي تُعطى بدلالة مُعاملات الارتباط المتعدِّد باستخدام العلاقة الآتية:

$$\mathbf{r}_{j,\ell;k} = 1 - \frac{1-\mathbf{r}^2_{j;\ell,k}}{1-\mathbf{r}^2_{j,\ell}} \qquad [3,52]$$

وأمّا إذا كان لدينا $n=4$ فعندئذ من أجل $j,\ell,k,m \in \mathbb{N}_4$ مع $j\neq\ell\neq k\neq m$ فإنَّ مُعاملات الارتباط الجزئي تُعطى بدلالة مُعاملات الارتباط المتعدِّد باستخدام العلاقة الآتية:

$$\mathbf{r}_{j,\ell;k,m}=\frac{\mathbf{r}_{j,\ell;k}-\mathbf{r}_{j,m;k}\cdot\mathbf{r}_{\ell,m;k}}{\sqrt{\left(1-\mathbf{r}^2_{j,m;k}\right)\cdot\left(1-\mathbf{r}^2_{\ell,m;k}\right)}} \qquad [3,53]$$

وهكذا بالنسبة إلى بقية الحالات الممكنة.

**(١, ٤, ٤, ٣) مثال**

لتكن لدينا البيانات الإحصائية الآتية:

الجدول (٣, ٢٥)

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $X_3$ | 11 | 10 | 9 | 10 | 9 | 10 | 11 | 12 |
| $X_3$ | 8 | 7 | 6 | 5 | 6 | 7 | 8 | 9 |
| $X_3$ | 9 | 8 | 7 | 6 | 5 | 4 | 3 | 2 |

فنجد أنَّ:

$$\mathbf{r}_{1,2}=0.843 \quad \& \quad \mathbf{r}_{1,3}=-0.394 \quad \& \quad \mathbf{r}_{2,3}=-0.356$$

ومن ثمَّ تكون قيمة مُعامل الارتباط الجزئي $\mathbf{r}_{1,2;3}$ تساوي:

$$\mathbf{r}_{1,2;3}=\frac{\mathbf{r}_{1,2}-\mathbf{r}_{1,3}\cdot\mathbf{r}_{2,3}}{\sqrt{\left(1-\mathbf{r}^2_{1,3}\right)\cdot\left(1-\mathbf{r}^2_{2,3}\right)}}=0.818$$

وأخيراً لنحسب مُعامل الارتباط المتعدِّد $\mathbf{r}_{1;23}$ بدلالة مُعاملات الارتباط الجزئية حيث لدينا:

$$1-\mathbf{r}^2_{1;2,3}=\left(1-\mathbf{r}^2_{1,2}\right)\cdot\left(1-\mathbf{r}^2_{1,3;2}\right)$$

علماً أنَّ:

$$\mathbf{r}_{1,3;2}=\frac{\mathbf{r}_{1,3}-\mathbf{r}_{1,2}\cdot\mathbf{r}_{2,3}}{\sqrt{\left(1-\mathbf{r}^2_{1,2}\right)\cdot\left(1-\mathbf{r}^2_{2,3}\right)}}=-0.187$$

وبالتعويض ينتج لدينا:

$$1-\mathbf{r}^2_{1;2,3}=\left(1-0.843^2\right)\cdot\left(1-(-0.187)^2\right)=0.279$$

ومن هذه الأخيرة نستنتج أنَّ $\mathbf{r}_{1;2,3}=0.721$ .

**هذا ما تيسَّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل**

# تمارين الفصل الثالث

١- لتكن لدينا مجموعة من البيانات الإحصائية التي تُمثِّل قياسات الجهد الكهربي والتيار $(x_i\ ,y_i)$ في 18 نقطة من دارة كهربائية:

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| ***X*** | 31 | 30 | 29 | 12 | 11 | 10 | 9 | 7 | 5 |
| ***Y*** | 35 | 34 | 33 | 15 | 16 | 16 | 15 | 14 | 13 |
| ***X*** | 30 | 20 | 19 | 18 | 17 | 20 | 19 | 18 | 17 |
| ***Y*** | 24 | 23 | 22 | 21 | 24 | 21 | 24 | 23 | 22 |

والمطلوب ما يلي:

أ - تمثيل هذه البيانات على لوحة الانتشار.

ب- احسب قيمة مُعامل ارتباط بيرسون لهذه المشاهدات.

ج- تعيين مستقيم انحدار $X$ على $Y$ وكذلك $Y$ على $X$ .

د- ارسم مستقيمي الانحدار السابقين على لوحة الانتشار.

هـ- استخدم مستقيم الانحدار المناسب لتقدير القيمة الموافقة لـ $x\ = 22.45$ وكذلك القيمة الموافقة لـ $y\ = 19.55$ .

٢- لتكن لدينا مجموعة من البيانات الآتية التي تعطينا $X$ مستوى تماسك مادة صناعية مع $Y$ درجات الحرارة لعشر عينات من هذه المادة:

| رقم العيِّنة | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| ***X*** | جيد | جيد جداً | وسط | جيد | وسط | جيد جداً | جيد | ضعيف | جيد | ضعيف جداً |
| ***Y*** | 9 | 10 | 6.5 | 7.5 | 8.5 | 9.5 | 7.5 | 4 | 8 | 1.5 |

والمطلوب تعيين قيمة مُعامل الارتباط لهذه البيانات، ومن ثَّم مناقشة النتيجة.

٣- لتكن لدينا مجموعة من البيانات الآتية التي تعطينا مستوى قدرات عشرة من الطلاب في مقرَّري الحوسبة $X$ والإحصاء $Y$:

| رقم العنصر | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| ***X*** | جيد | جيد جداً | وسط | جيد | وسط | جيد جداً | جيد | ضعيف | جيد | ضعيف جداً |
| ***Y*** | جيد جداً | جيد جداً | وسط | وسط | جيد | جيد | جيد | وسط | جيد | ضعيف |

والمطلوب تعيين قيمة مُعامل الارتباط لهذه البيانات، ومن ثَّم مناقشة النتيجة.

٤- لتكن لدينا بيانات إحصائية معطاة من خلال الجدول الآتي:

| | | | رتب $X$ | | | | | Total |
|---|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 | |
| | | حدود الفئة | 5→10 | 10→15 | 15→20 | 20→25 | 25→30 | |
| رتب $Y$ | 1 | 5→15 | 15 | 5 | - | 1 | - | |
| | 2 | 15→25 | 20 | 13 | 5 | 2 | 1 | |
| | 3 | 25→35 | 30 | 28 | 16 | 10 | 5 | |
| | 4 | 35→45 | 15 | 12 | 5 | - | 1 | |
| | 5 | 45→55 | 6 | 4 | 2 | - | - | |
| Total | | | | | | | | |

والمطلوب تعيين قيمة مُعامل الارتباط لهذه البيانات، ومن ثَمَّ مناقشة النتيجة.

٥- لتكن لدينا مجموعة من البيانات الإحصائية الآتية:

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $X$ | 0.2 | 0.35 | 1 | 0.25 | 1.1 | 3 | 2.5 | 1.75 | 1.5 | 0.75 |
| $Y$ | -1.61 | -1.05 | 0.1 | -1.5 | 0.1 | 1.15 | 0.95 | 0.57 | 0.35 | - 0.3 |

والمطلوب ما يلي:

أ- تمثيل هذه البيانات على لوحة الانتشار.

ب- تعيين خط الانحدار المناسب لهذه البيانات.

جـ- تعيين معادلة هذا الخط بعد ردها إلى الشكل الخطي.

٦- لتكن لدينا مجموعة من البيانات الإحصائية الآتية:

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $X$ | 2 | 0.35 | 1.5 | 2 | 2.75 | 3 | 2.5 | 3.5 | 4.5 | 0.75 |
| $Y$ | 4.5 | 0.1 | 2.2 | 3.75 | 7.7 | 9.5 | 6 | 12.5 | 21 | 0.5 |

والمطلوب ما يلي:

أ- تمثيل هذه البيانات على لوحة الانتشار.

ب- تعيين خط الانحدار المناسب لهذه البيانات.

جـ- تعيين معادلة هذا الخط بعد ردها إلى الشكل الخطي.

٧- لتكن لدينا مجموعة من البيانات الإحصائية الآتية:

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $X$ | 1 | 3 | 0.5 | 0.25 | 1.1 | 2 | 2.5 | 1.75 | 1.25 | 2.75 |
| $Y$ | 1.2 | 0.4 | 2.1 | 4.1 | 0.9 | 0.47 | 0.45 | 0.56 | 0.7 | 0.38 |

والمطلوب ما يلي:

أ- تمثيل هذه البيانات على لوحة الانتشار.

ب- تعيين خط الانحدار المناسب لهذه البيانات.

جـ- تعيين معادلة هذا الخط بعد ردها إلى الشكل الخطي.

٨- لتكن لدينا مجموعة من البيانات الإحصائية الآتية:

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $X$ | 0.2 | 1.35 | 1.75 | 2.25 | 1.5 | 3 | 2.5 | 0.8 | 1.5 | 0.5 |
| $Y$ | 1.3 | 3.75 | 5.8 | 10 | 4.6 | 20 | 12.3 | 2.1 | 5.1 | 1.7 |

والمطلوب ما يلي:

أ- تمثيل هذه البيانات على لوحة الانتشار.

ب- تعيين خط الانحدار المناسب لهذه البيانات، ومن ثَمَّ تعيين معادلة هذا الخط بعد ردها إلى الشكل الخطي.

٩- في دراسة لمعرفة مدى تأثير النبضات العابرة المتراكبة على خط الجهد المغذي للمنازل 220 فولط (هذه النبضات تنتج من بدء إقلاع الأجهزة الموصولة على هذه الخطوط وقد تبلغ نحو 600 فولت أو أكثر ولكن بتردد عال نسبياً) على نوع من الأجهزة الإلكترونية الدقيقة أخضعت عيِّنة عشوائية مكوَّنة من خمسين جهازاً لهذه الدراسة، وبحيث رُبط منها عشرون جهازاً على خط تغذية مستقل عن الخط العامَ للتغذية من أجل عمليّة المقارنة، فوجدنا النتائج الآتية:

| | | الظاهرة الأولى (التعريض للنبضات) | | Total |
|---|---|---|---|---|
| | | أجهزة وصلت إلى خـط التغذية العـام (تَعرَّضت لنبضات عابرة) | أجهزة وصلت إلى خط تغذية مستقل (لم تتعرض لنبضات عابرة) | |
| الظاهرة الثانية (الإصابة بالتلف) | أصيبت بالتلف | 4 | 1 | |
| | لم تصب بالتلف | 26 | 19 | |
| Total | | | | |

والمطلوب تبيان العلاقة ما بين ظاهرة تلف الأجهزة وظاهرة التعرض للنبضات العابرة المحمولة على خط التغذية العام.

١٠- لتكن لدينا البيانات الإحصائية الآتية:

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $X_1$ | 10 | 9 | 8 | 7 | 4 | 5 | 10 | 7 |
| $X_2$ | 9 | 8 | 7 | 6 | 5 | 6 | 9 | 8 |
| $X_3$ | 8 | 7 | 6 | 5 | 6 | 7 | 8 | 9 |

والمطلوب ما يلي:

أ- حساب قيم مُعاملات الارتباط المتعدِّد $\mathbf{r}_{1;23}$ و $\mathbf{r}_{3;12}$ بدلالـة مُعاملات الارتباط البسيط.

ب- حساب قيم مُعاملات الارتباط الجزئية $\mathbf{r}_{1,2;3}$ و $\mathbf{r}_{1,3;2}$ بدلالـة مُعاملات الارتباط البسيط.

١١- لدى إجراء دراســة على مدى تأثير الجهد الكهربائي ذي التوتر العالي (والمُحمَّل على خطوط التغذية القريبة من المنازل الســكنية) على صحة الأشخاص القاطنين على مسافة عشرة أمتار عن هذه الخطوط وجدت النتائج المدوَّنة كما في الجدول الآتي:

| | | الظاهرة الأولى $X$ (التوتّر) | | | | Total |
|---|---|---|---|---|---|---|
| | | توتر منخفض $220\ V$ | توتر متوسط $5000\ V$ | توتر عـال $20\ K\ V$ | توتر عال جداً $100\ KV$ | |
| الظاهرة الثانية $Y$ (الحالة الصحية) | أشخاص لم يعانوا من أمراض | 100 | 85 | 15 | 5 | |
| | أشخاص بدت عليهم أمراض | 5 | 15 | 25 | 50 | |
| Total | | | | | | |

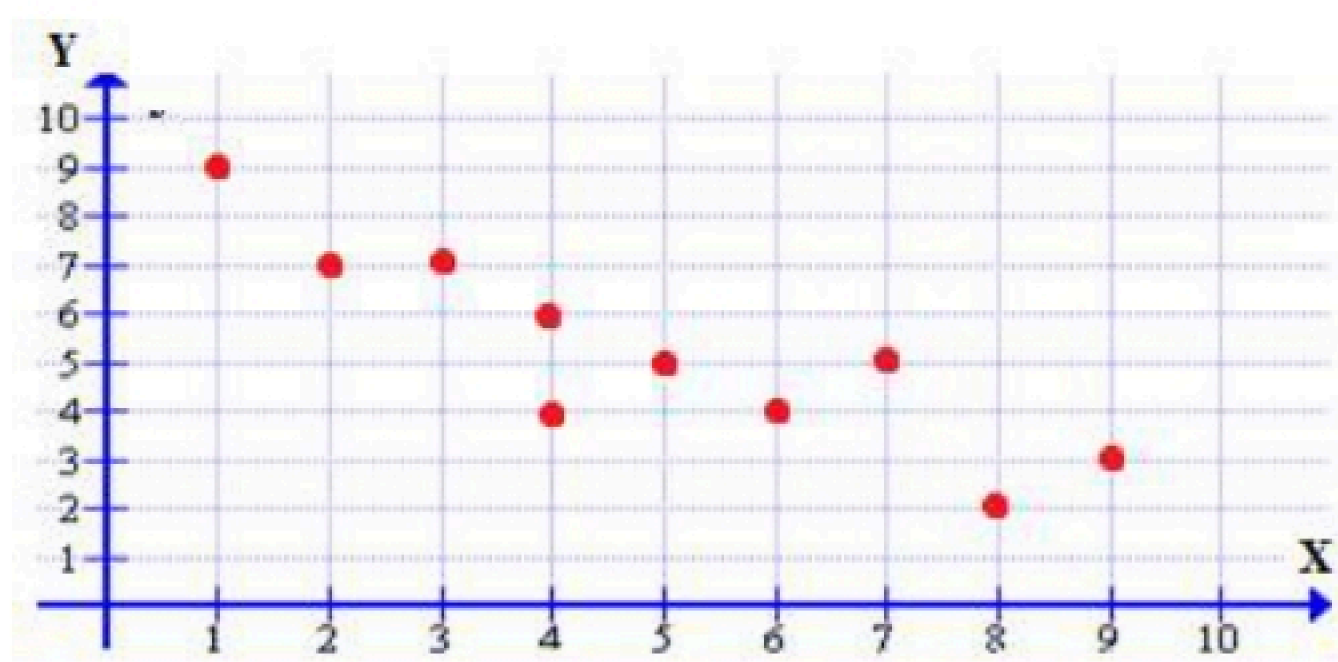
والمطلوب تبيان مدى ارتباط ظاهرة الأمراض التي بدت على الأشخاص مع قيمة التوتر الكهربائي القريب من موقع سكنهم.

**١٢-** لتكن لدينا مجموعة بيانات إحصائية $(x_i\,, y_i)$ مع $i \in \mathbb{N}_{10}$ ومقدّمة من خلال لوحة الانتشار الآتية، والمطلوب ما يلي:

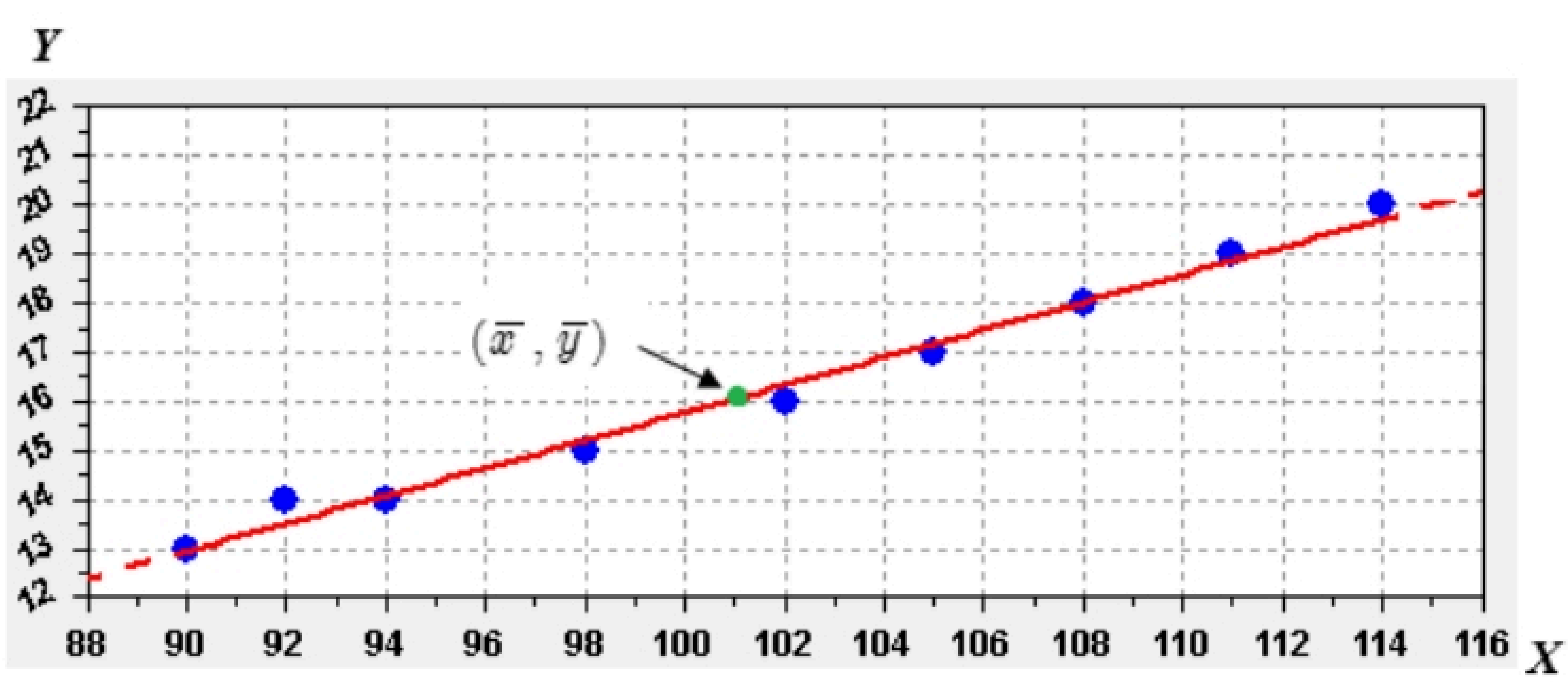
أ- حساب قيمة مُعامل ارتباط بيرسون لهذه المشاهدات.

ب- تعيين معادلة مستقيم انحدار $X$ على $Y$، ومن ثَمَّ رسمه على لوحة الانتشار.

ج- استخدم معادلة مستقيم الانحدار لتقدير القيمة الموافقة لـ $x = 3.5$ .




**١٣-** لتكن لدينا مجموعة بيانات إحصائية $(x_i\,, y_i)$ مع $i \in \mathbb{N}_{10}$ ومقدّمة من خلال لوحة الانتشار الآتية




والمطلوب ما يلي:

أ- حساب قيمة مُعامل ارتباط بيرسون لهذه المشاهدات.

ب- استنتاج معادلة مستقيم انحدار $X$ على $Y$ من الرسم المُعطى.

ج- استخدم معادلة مستقيم الانحدار لتقدير القيمة الموافقة لـ $x = 104$ .

# الفصل الرابع

## الفضاء الاحتمالي، الاحتمالات الشرطية واستقلال الحوادث

**PROBABILITY SPACE, CONDITIONAL PRPBABILITY & INDEPENDENCE OF EVENTS**

### (٤,١) الفضاء الاحتمالي لتجربة عشوائية

**Probability Space for a Random Experiment**

إنَّ مفهوم الفضاء الاحتمالي لا يُعدّ من المفاهيم الرياضياتية القديمة إذا ما قورن بمفاهيم رياضياتية أخرى مثل الفضاء المتَّجهي وفضاء هيلبرت، حيث تبلور هذا المفهوم بشكل واضح بعد تقديم ما يُعرف باسم **مسلَّمات كلموغوراف** للفضاء الاحتمالي (**عام 1933**) التي قُدِّمت من قبل الرياضياتي الروسي **كلموغوراف** (1903-1987) Andrey Nikolaevich Kolmogorov حيث أُخذ هذا التاريخ كنقطة بداية للنظرية الحديثة للاحتمالات، وذلك لأنَّ هذه المسلَّمات لم تستوعب قطاع الاحتمالات التقليدية وحسب، وإنَّما استطاعت أن تنسجم مع متطلّبات العلوم الطبيعية والتقنية لدى دراسة التوزيعات ذات ا لبُعد غير المنته أيضاً.

في الحقيقة إنَّ أوَّل بحث قُدِّم في إطار النظرية الحديثة للاحتمالات يعود إلى الرياضياتي الروسي **برنشتاين** Sergei Natanovich Bernstein (1880-1968)، ولّلرياضياتي الفرنسي **بوريل** (1871-1956) Félix Édouard Justin Émile Borel، إلاَّ أنَّ كلموغوراف كان له الأثر الأكبر في تلك الآونة، وذلك بعد أن قام بنشر كتابه **المفاهيم الأساس للحساب الاحتمالي** عام 1933 الذي تضمَّن مسلَّمات الحساب الاحتمالي، والمسلَّمات العامّة لنظرية الاحتمالات القائمة على نظرية القياس. كما أُدخل مفهوم جبر الحوادث لأوَّل مرَّة من قبل الرياضياتي والفيلسوف الأكراني **غليفينكو** (1897-1940) Valery Ivanovich Glivenko وذلك عام 1939.

الآن، ولتوضيح مفهوم الفضاء الاحتمالي لا بدَّ لنا أولاً أن نتعرّف على ما تعنيه عبارة علم الاحتمالات، وهذا المصطلح ذو علاقة وطيدة بمفهوم التجربة العشوائية، علماً أنَّ الإحصائيين يستخدمون كلمة تجربة لوصف أي عمليّة تولّد مجموعة من البيانات، ومن أجل الوصول إلى هذا التوضيح سننطلق من الحقيقة الآتية:

إنَّ التجارب التي يقوم بها المرء تقسم عادةً إلى صنفين:

**الأول:** تجارب تكون نتائجها معروفة مسبقاً وقابلة للتخمين بشكل دقيق، ومن الأمثلة علـى ذلك تجربة تفاعل حـمض كلور الماء مع كمية من الألمنيوم النقي، فعندئذ سينتج لدينا (**وبكل تأكيد**) كلور الألمنيوم وهيدروجين بالإضافة إلى كمية من الحـرارة، وأكثر من ذلك يمكننا حساب الكميات الناتجة بدقة عالية جداً عند معرفة الكميات الداخلة في التفاعل. إنَّ هذا الصنف من التجـارب التي يمكن الحكم على نتائجها مسبقاً وبدقة تُدعى **تجارب نظامية** Regular Experiments.

**الثاني:** خاص بالتجارب التي لا يمكن معرفة نتائجها بشكل دقيق مسبقاً، وكل ما يُعرف عنها أنَّ النتائج ستنتمي إلى مجموعة ما يمكن تعيينها، ومثال على ذلك تجربة إلقاء حجر زهر النرد Die سداسي الوجوه (**يوجد**

أنواع عديدة من أحجار النرد، ولكن القياسي منها على شكل مكعّب)، فعندئذ ستكون النتيجة هي ظهور أحد الأرقام 1 وحتى 6. أي إنَّ النتيجة التي ستظهر للأعلى لا يمكن معرفتها مسبقاً أو الجزم بحدوثها، ولكن ستنتمي للمجموعة $\{1,2,3,4,5,6\}$ بكل تأكيد. إنَّ هذا الصنف من التجارب يُدعى **تجارب عشوائية** Random Experiments.

في هذا الموضع قد يتبادر إلى ذهن القارئ طرح السؤال الآتي:

**ما هي العشوائية؟**

في الواقع إنَّ الإجابة عن هذا السؤال ترتكز على أساسين:

**الأول**: رياضياتي بَحت، ويهتمّ بما يعنيه التتالي العشوائي للقيم العددية أو الدوال أو القياسات أو ...

**الثاني**: تطبيقي صِرف، ويهتمّ بالمسألة التي تُحدِّد إن كانت الحوادث التي تجري في الواقع خاضعة للمعايير الرياضياتية الخاصة بالعشوائية أم لا.

هكذا نلاحظ أنَّه لا يمكننا الجزم بتتابع حوادث تطبيقية المنشأ، أهو تتابع عشوائي أم لا؟ قبل وضع تعريف رياضياتي دقيق لمعنى العشوائية، وهنا نلاحظ أنَّه حتى في حال تقديم مثل هذا التعريف، فإنَّ المشكلة تبقى قائمة بخصوص توافق الحوادث التطبيقية مع هذا التعريف. إذ إنَّه من الممكن أن تتطابق العلاقات الجبرية الممثِّلة لنتائج التجربة مع القانون المولِّد للقيم العشوائية التي سيتم اعتمادها في الدراسة، وفي مثل هذه الحالة ستزول ظاهرة العشوائية، وستبدو نتائج الدراسة على أنَّها قيم نظامية وليست عشوائية.

**إذن ما العمل لحل هذه المعضلة؟**

في الواقع لم يُقدَّم حتى اليوم تعريف دقيق شامل لمعنى العشوائية رغم الاستخدام الهائل لهذا المصطلح والعدد الكبير جداً للكتب المؤلَّفة في هذا الصدد، وربما يكون السبب في ذلك (وهذا رأي شخصي) هو العدد الكبير (لا بل الكبير جداً) من العوامل التي تؤثِّر في نشوء ظاهرة العشوائية، ومنها على سبيل المثال لا الحصر:

١ – الأخطاء النّاتجة عن القياس، وذلك لأنَّ أي جهاز قياس (ومهما بلغت دقته) سيكون له ارتياب ولو بقدرٍ ضئيلٍ جداً، وهذا الارتياب قد يمكن تقديره، ولكن لا يمكن الجزم بقيمته.

٢ – المؤثِّرات الخارجية في مادة التجربة التي لا يمكن تفاديها أو السيطرة عليها، وهذه المؤثّرات كثيرة جداً قد تصل إلى حد تشكيل طيف من المؤثرات المعلوم بعضها، والمجهول بعضها الآخر (كما هو الحال لدى التنبؤ بهطول الأمطار مثلاً).

٣– مهارة المجرِّب في تنفيذ التجربة، ويضاف إليها الحالة النفسية للمجرِّب عند لحظة تنفيذ التجربة (كأن يتلقّى نبأً غير سار مثلاً).

إذن والحال كذلك كان لا بدَّ من اتفاق على معنى العشوائية. إنَّ هذا الاتفاق ما زال موضع جدل. فمنهم من يعزوها إلى الصدفة البحتة (وهل هناك صدفة بحتة وأخرى غير بحتة، فإن كان هناك مثل هذا التصنيف فإنَّه من المنطقي أن تنفي إحداهما الأخرى)، وهنا يُطرح الرأي القائل:

ما هو صدفة بالنسبة لك قد يكون رتيباً بالنسبة لآخر، والأمثلة على ذلك كثيرة جداً، فعلى سبيل المثال: يقوم رجل عدواني الطبع والمزاج بالاعتداء على رجل آخر مجهول بالنسبة له بقصد السلب، فإنَّ ما حصل للرجل المعتدى عليه هو صدفة محضة، في حين أنَّ حدوث الاعتداء كان رتيباً ومخطَّطاً بالنسبة للرجل المعتدي.

في دراستنا هذه سنقبل بأنَّ العشوائية هي مجموعة العوامل والمؤثِّرات التي تفرض نفسها على التجربة بحيث يصبح معها إمكانية تحديد النتيجة بدقّة مطلقة مسبقاً أمر غير ممكن. هذا وقد وردت عبارة العشوائية كترجمة للكلمة اللاتينية Stochastic وتعني يصيب بسهم، وهذه الكلمة استخدمت فيما بعد لتعني **علم العشوائيات** أيضاً الذي هو أحد فروع العلوم الرياضياتية، ويقصد به عادةً **نظرية الاحتمالات** Probability Theory **والإحصاء الرياضياتي** Mathematical Statistics (انظر تاريخ الاحتمالات [51]).

الآن، وبما أنَّ المرء يهتمّ عادة بالنمذجة الرياضياتية للتجارب، فإنَّه من الأهمية بمكان معرفة العلم الذي يساعده في عملية النمذجة لهذه التجارب، فعلى سبيل المثال نجد أنَّ النمذجة الرياضياتية للتجربة النظامية السابقة تُقدَّم من خلال المعادلة الرياضياتية الآتية:

$$2Al + 6HCl = 3H_2 + 2AlCl_3$$

وهي تندرج في إطار علم الحساب (وهناك بعض التجارب النظامية الأكثر تعقيداً قد تحتاج إلى نماذج رياضياتية معقدّة تتطلب علم التحليل الرياضياتي والجبر).

أمّا بخصوص التجارب العشوائية، فإنَّ العلم الذي يهتمّ بتقديم النماذج الرياضياتية النظرية لها هو علم الاحتمالات. بمعنى أنَّ علم الاحتمالات هو ذلك الفرع من الرياضيات الذي من ضمن اهتماماته إعداد النماذج الرياضياتية النظرية للتجارب العشوائية التي تُعرف باسم الفضاءات الاحتمالية للتجارب العشوائية أيضاً، وهذا العلم يُعدُّ بمثابة حجر الأساس لعلم العشوائيات.

سنقدِّم فيما يلي عرضاً لكيفية تعيين الفضاء الاحتمالي (النموذج الرياضياتي النظري) لتجربة عشوائية، ولكن على سبيل التوضيح سنقوم أولاً بتقديم الفضاء الاحتمالي من أجل تجارب عشوائية ذات عدد منته من النتائج، وبعد ذلك نقدِّم الفضاء الاحتمالي بشكله العام بناءً على مسلَّمات حيث سيتضّح لنا سبب هذا التعميم من خلال مجرّيات هذا الفصل. لكن نودّ أن نلفت انتباه القارئ إلى أنَّ الفهم الصحيح والدقيق للاحتمالات يتطلب معارف من نظرية القياس Measure Theory، ولذلك قمنا بتقديم موجز لأهم المعلومات من نظرية القياس في الملحق (**B**) قد نحتاجها في هذا الكتاب. لهذا يجب على القارئ الاطلاع على ذلك الملحق إن كان غير ملمّ بهذه النظرية لأنّنا سنردّ بعض المفاهيم هنا إلى المعلومات المقدَّمة في ذلك الملحق. أمّا المفاهيم الرياضياتية التي سنحتاج إلى بعضها فيفترض أن يكون القارئ ملمَّا بها، وعلى سبيل التذكير قمنا بتخصيص الملحق (**A**) لتدوين ما هو ضروري منها.

من أجل تجربة عشوائية مفترضة يُرمز لفضائها الاحتمالي بثلاثية $[\Omega, \mathscr{A}, P]$ ، وهكذا نلاحظ أنَّ تعيين الفضاء الاحتمالي لأية تجربة عشوائية يحتاج إلى تعيين ثلاثة عناصر رئيسة على التوالي هي:

$\Omega$ وهي مجموعة كل نتائج التجربة العشوائية.

$\mathscr{A}$ وهي أسرة لمجموعات جزئية من $\Omega$ تحقّق شروطاً محدَّدة (سنأتي على ذكرها لاحقاً)، ويُدعى جبر الحوادث Algebra of Events للتجربة العشوائية.

$P$ وهي دالَّة حقيقية معرَّفة على أسرة المجموعات $\mathscr{A}$ ، وتحقِّق شروطاً محدَّدة (سنأتي على ذكرها لاحقاً)، وتُدعى قياساً احتمالياً Probability Measure (أو دالَّة احتمالية).

الآن لشرح وتوضيح كل عنصر من العناصر الثلاث السابقة، سنفترض أنّنا أمام تجربة عشوائية مجموعة نتائجها الممكنة $\Omega$ (في إطار دراستنا المقبلة سنفترضها غير خالية دائماً).

### (١ , ١ , ٤) تعيين مجموعة نتائج تجربة عشوائية Determination the Set of Outcomes of a Random Experiment

من أجل توضيح كيفية تعيين مجموعة نتائج تجربة عشوائية (أو فضاء الحوادث الابتدائية) سنقدّم نوعين من الأمثلة. الأول منها يتعلّق بتجارب عشوائية ذات عدد منته من النتائج، وأمّا الآخر فإنَّه يتعلّق بتجارب عشوائية ذات عدد غير منته من النتائج، ولكن قبل البدء بتقديم هذه الأمثلة نودّ تقديم الملاحظات الهامَّة الآتية:

١ – عندما نذكر مستقبلاً أنَّه في تجربة عشوائية ما (مثل: قذف قطع نقود معدنية، أو رمي حجارة نرد، أو سحب كرات من صندوق، أو...) أنَّ الأدوات المستخدمة في توليد النتائج (مثل: قطع النقود المعدنية، أو حجارة النرد، أو الكرات التي في صندوق، أو...) هي أدوات متماثلة إنَّما نقصد بذلك أنَّه لا يمكننا التمييز بين هذه الأدوات (انظر الفقرتين /٤/ و /٧/ من المثال الآتي)، وفي حال ذكرنا أنَّ هذه الأدوات متمايزة إنَّما نقصد بذلك أنَّ النتائج المولَّدة بتلك الأدوات يمكن تمييز بعضها عن البعض الآخر (انظر الفقرتين /٣/ و /٦/ من المثال الآتي).

٢- عندما نذكر مستقبلاً أنَّ الأدوات المستخدمة في التجربة (مثل: قطع نقود، أو حجارة نرد، أو.....) متوازنة، فإنَّنا نقصد بذلك أنَّ هذه الأدوات مصنَّعة من مادة متجانسة الكثافة بحيث يصبح لكل نتيجة من نتائج التجربة العشوائية النصيب نفسه في الظهور.

### (١ , ١ , ١ , ٤) أمثلة على تجارب ذات عدد منته من النتائج

في الأمثلة الآتية لن نهتمّ بطبيعة النتّائج التي سنحصل عليها، وإنَّما المهمّ هنا أن يكون عدد هذه النتائج منته

١- لنأخذ تجربة قذف قطعة نقود Coin (ويُقصد بذلك قطعة نقود معدنية دوماً) لمرّة واحدة فقط، فعندئذ نحصل على إحدى نتيجتين، إمّا صورة Head ويُرمز لها بـ$H$، أو شعار Tail ويُرمز لها بـ$T$. إنَّ سبب استخدام Head تعود إلى بعض القطع النقدية المعدنية التي استخدمت في دراسة هذا النوع في التجارب قديماً، حيث كان ينقش على أحد وجهيها صورة رأس أو جذع مع الرأس لشخصية شهيرة في البلد، وأمّا على الوجه الآخر فإنَّه ينقش شعار البلد.

الشعار $T$ الصورة $H$ الشعار الكتابة عوضاً عن الصورة

بعض النماذج من العملة المعدنية

وهكذا نجد أنَّ مجموعة نتائج هذه التجربة العشوائية هي $\Omega = \{H, T\}$، حيث نلاحظ أنَّ عدد النتائج الممكنة لهذه التجربة يساوي 2، وهذا يمثّل قدرة مجموعة نتائج التجربة العشوائية $\Omega$، أي إنَّ $|\Omega| = 2$، علماً أنَّ $|\Omega|$ يرمز إلى قدرة (عدد عناصر) المجموعة $\Omega$.

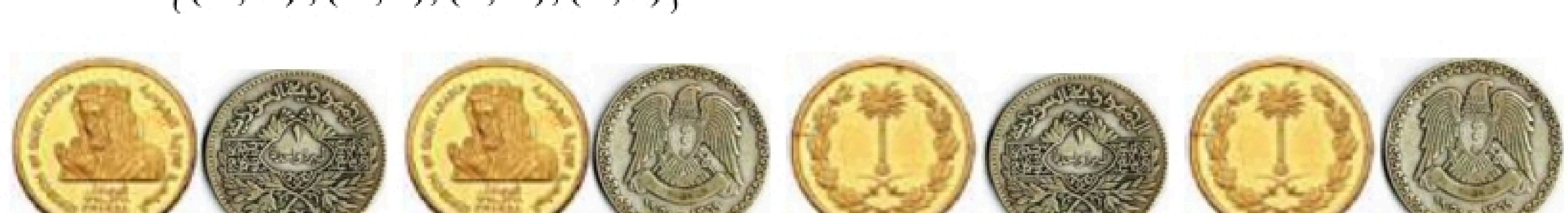
٢- لنأخذ تجربة قذف قطعة نقود واحدة لمرّتين متتاليتين، فعندئذ ستكون نتائج التجربة هي عبارة عن ثنائيات مركباتها إمّا شعاراً أو صورةً، ومن ثمَّ فإنَّ مجموعة النتائج لهذه التجربة العشوائية هي:

$$\Omega = \{(H,H), (H,T), (T,H), (T,T)\}$$

حيث نلاحظ أنَّ عدد النتائج الممكنة لهذه التجربة يساوي $|\Omega| = 4 = 2^2$.

٣- لنأخذ تجربة قذف قطعتي نقود متمايزتين لمرّة واحدة فقط، فعندئذ ستكون نتائج التجربة هي عبارة عن ثنائيات أيضاً، ومركباتها إمّا شعار أو صورة، ومن ثمَّ فإنَّ مجموعة النتائج لهذه التجربة العشوائية هي:

$$\Omega = \{(H,H), (H,T), (T,H), (T,T)\}$$



وهنا نلاحظ أنَّ مجموعة النتائج الممكنة لهذه التجربة متطابقة مع مجموعة النتائج الممكنة للتجربة في المثال السابق، والسبب في ذلك يعود لتمايز قطعتي النقود.

ننوِّه هنا إلى أنَّنا سنكتب (في معظم الحالات- وعلى سبيل التبسيط-) المجموعة السابقة $\Omega$ كما يلي:

$$\Omega = \{ HH , HT , TH , TT \}$$

إذا كان ذلك لا يؤدي إلى التباس، وبالمثل بالنسبة إلى المجموعات التي تماثلها.

٤- لنأخذ تجربة قذف قطعتي نقود غير متمايزتين (متماثلتين تماماً) لمرّة واحدة فقط، فعندئذ ستكون نتائج التجربة هي:

وكما هو ملاحظ النتائج عبارة عن ثنائيات أيضاً، ولكنَّ مجموعة النتائج الممكنة لهذه التجربة هي:

$$\Omega = \{ (H , H) , (H , T) , (T , T) \}$$

ومن ثمَّ نلاحظ أنَّ عدد النتائج الممكنة لهذه التجربة تساوي $|\Omega| = 3$، وذلك لأنَّه بسبب عدم التمايز لقطعتي النقود فإنَّه يُنظر إلى النتيجتين:

$(H,T)$   $(T,H)$

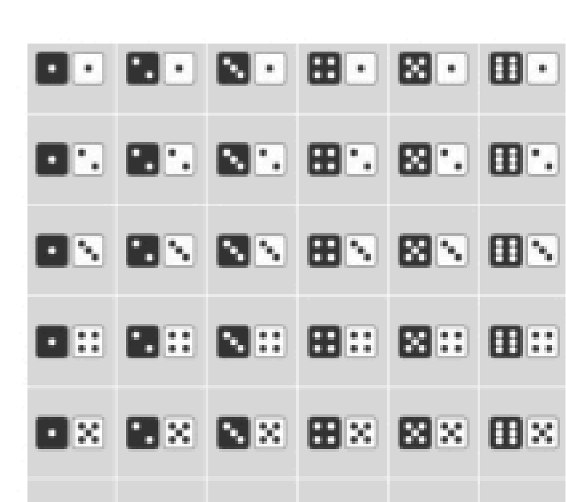
على أنَّها نتيجة واحدة فقط حيث لا يمكننا معرفة تبعية المركَّبة الأولى (وبالمثل المركَّبة الثانية) أهي من القطعة الأولى أم الثانية؟

٥- لنأخذ تجربة رمي حجر نرد Die لمرّة واحدة فقط، فعندئذ ستكون مجموعة النتائج لهذه التجربة العشوائية هي:

$$\Omega = \{1 , 2 , 3 , 4 , 5 , 6 \}$$

ومن ثمَّ يكون لدينا $|\Omega| = 6$.

٦- لنأخذ تجربة رمي حجري نرد Two Dice متمايزين ولمرّة واحدة فقط، فعندئذ ستكون مجموعة النتائج لهذه التجربة العشوائية هي:



$$\Omega = \left\{ \begin{matrix} (1,1),(1,2),(1,3),(1,4),(1,5),(1,6), \\ (2,1),(2,2),(2,3),(2,4),(2,5),(2,6), \\ (3,1),(3,2),(3,3),(3,4),(3,5),(3,6), \\ (4,1),(4,2),(4,3),(4,4),(4,5),(4,6), \\ (5,1),(5,2),(5,3),(5,4),(5,5),(5,6), \\ (6,1),(6,2),(6,3),(6,4),(6,5),(6,6) \end{matrix} \right\}$$

ومن ثمَّ يكون لدينا $|\Omega| = 36 = 6^2$.

٧- لنأخذ تجربة رمي حجري نرد غير متمايزين (متماثلين تماماً) ولمرّة واحدة فقط، فعندئذ بسبب التماثل بين حجري النرد سوف يُنظر إلى النتيجتين $(1,2)$  و$(2,1)$  على أنَّها تُمثِّل نتيجة واحدة فقط (ولكن بإمكَانية الظهور لمرتين) وذلك لأنَّنا لا نستطيع البت في مصدر العدد 2 (وبالمثل بالنسبة إلى مصدر العدد 1) من أي الحجرين قد نتج، وبناء على ذلك تُؤخذ إحداهما كممثِّل للنتيجة فقط. بالمثل بالنسبة لبقية المركَّبات $(i,j)$ و $(j,i)$ مع $i \neq j$ و $i,j \in \mathbb{N}_6$، ومنه يكون لمجموعة نتائج هذه التجربة العشوائية العرض الآتي:

$$\Omega=\begin{Bmatrix}(1,1),(1,2),(1,3),(1,4),(1,5),(1,6),\\(2,2),(2,3),(2,4),(2,5),(2,6),\\(3,3),(3,4),(3,5),(3,6),\\(4,4),(4,5),(4,6),\\(5,5),(5,6),\\(6,6)\end{Bmatrix}$$

ومن ثَّم يكون لدينا $|\,\Omega\,|=21$ .

٨- لنأخذ تجربة قذف قطعة نقود لـ $n$ مرة متتالية، فعندئذ تكون مجموعة النتائج الممكنة لهذه التجربة العشوائية هي متتالية منتهية من رموز الصور $H$ والشعار $T$ متوضِّعة بترتيب معيَّن، ومن ثَّم يكون لمجموعة النتائج لهذه التجربة العشوائية العرض الآتي:

$$\Omega=\{\,\omega:\omega=(a_1,a_2,\ldots,a_n)\;;\;a_i\in\{H,T\}\;\&\;i\in\mathbb{N}_n\,\}$$

فنلاحظ أنَّ $|\,\Omega\,|=2^n$ ، كما يمكننا الحصول على مجموعة النتائج السابقة نفسها أيضاً من خلال رمي $n$ قطعة نقود متمايزة في آن واحد ولمرَّة واحدة فقط.

٩- لنأخذ تجربة إلقاء حجري نرد متمايزين في آن واحد ولـ $n$ مرّة متتالية، فعندئذ يكون لمجموعة النتائج الممكنة لهذه التجربة العشوائية العرض الآتي:

$$\Omega=\{\,\omega:\omega=(a_1,a_2,\ldots,a_n)\;;\;a_i=(b_1,b_2)\;\&\;b_1,b_2\in\mathbb{N}_6\;\&\;i\in\mathbb{N}_n\,\}$$

فنلاحظ أنَّ عدد النتائج الممكنة لهذه التجربة يساوي $\left(6^2\right)^n=36^n=|\,\Omega\,|$ .

### ١٠ - تجربة السحب مع الإعادة Pulling out with Replacement

لنفترض أنَّه لدينا وعاء يحوي $m$ كرة مرقمة من 1 وحتى $m$، وسنفترض أنَّها تتمايز بعضها عن بعض من خلال الرقم المدوَّن عليها فقط. تُخلط هذه الكرات جيداً ومن ثَّم تُسحب كرة من الوعاء (دون التعمّد في سحبها، أي بطريقة عشوائية) ويُسجَّل رقمها، ثَّم نعيدها إلى الوعاء ثانيةً. تُخلط الكرات من جديد بشكل جيد ومن ثَّم تُسحب كرة من الوعاء (دون التعمّد في سحبها) ويُسجَّل رقمها ثَّم تعاد إلى الوعاء من جديد، وعلى هذا النحو تُكرِّر هذه العملية لـ $n$ مرَّة متتالية (إنَّ هذا النوع من التجارب يعرف باسم تجربة السحب مع الإعادة)، فإذا كنّا مهتمين بترتيب نتائج السحب، أي إنَّنا مهتمين بترتيب النتائج كما نحصل عليها وبغض النظر عن حصولنا عليها سابقاً أم لا، فعلى سبيل المثال يُنظر إلى تسلسل النتائج الآتية على أنَّها نتائج مختلفةً بعضها عن البعض الآخر:

$$\underbrace{1,2,3,2,4,5,6,\ldots,k}_{n\text{- نتيجة}}\;\&\;\underbrace{1,2,2,3,4,5,6,\ldots,k}_{n\text{- نتيجة}}\;\&\;\underbrace{1,2,3,4,2,5,6,\ldots,k}_{n\text{- نتيجة}}\;\&\ldots$$

مع $1\le k\le m$، فعندئذ سيكون لمجموعة نتائج هذه التجربة العشوائية العرض الآتي:

$$\Omega=\{\,\omega:\omega=(a_1,a_2,\ldots,a_n)\;;\;a_i\in\mathbb{N}_m\;\&\;i\in\mathbb{N}_n\}\quad\Rightarrow|\,\Omega\,|=m^n$$

وأمّا إذا كنّا غير مهتمين بترتيب نتائج السحب (بمعنى أنّنا لا نهتمّ بترتيب المركّبات) ومن ثَّم لا نهتمّ في أي سحب سنحصل على رقم معيَّن، وإنَّما ما يهمُّنا هو إن كنَّا سنحصل على هذا الرقم أم لا، فعلى سبيل المثال يُنظر إلى تسلسل النتائج الآتية على أنَّها تمثِّل نتيجة واحدة فقط:

$$\underbrace{1,2,3,2,4,5,6,\ldots,k}_{n\text{- نتيجة}}\;\&\;\underbrace{1,2,2,3,4,5,6,\ldots,k}_{n\text{- نتيجة}}\;\&\;\underbrace{1,2,3,4,2,5,6,\ldots,k}_{n\text{- نتيجة}}\;\&\ldots$$

مع $1 \le k \le m$ ، فعندئذ سيكون لمجموعة نتائج هذه التجربة العشوائية العرض الآتي:

$$\Omega = \{ \omega : \omega = [a_1 , a_2 ,..., a_n ] \ ; a_i \in \mathbb{N}_m \ \& \ i \in \mathbb{N}_n \} \quad \Rightarrow \ |\Omega| = \binom{m+n-1}{n}$$

علماً أنَّ القوس المتوسط المستخدم هنا نعني به أنَّه لا أهمية للترتيب بين المركّبات (انظر الجدول (١ , ٤)).

### ١١ - تجربة السحب بدون إعادة Pulling out without Replacement

في كثير من هذه المجالات يتم الاختبار على حساب العنصر الذي خضع للاختبار، وهذا يعني أنَّ الكائن الذي يتم اختياره قد يتمّ إتلافه نتيجة الاختبار، ومن ثَمّ لا يمكن استبداله بآخر لتنفيذ الاختبار مرّة أخرى، ولذلك فإنَّ التعامل مع السحب دون إعادة (أو دون إحلال) ضروري في مثل هذه الحالات.

لنفترض أنَّ لدينا وعاء يحوي $m$ كرة مرقمة من 1 وحتى $m$، ويتمايز بعضها عن بعض بالرقم المدَّون عليها فقط. تُخلط هذه الكرات جيداً ومن ثمَّ تُسحب كرة من الوعاء (دون التعمّد في سحبها) ويُسجَّل رقمها، ومن ثمَّ توضع جانباً خارج الوعاء، وبعد ذلك تُخلط الكرات المتبقية من جديد بشكل جيد ومن ثمَّ تُسحب كرة من الوعاء (دون التعمّد في سحبها ) ويُسجَّل رقمها وبعد ذلك توضع جانباً خارج الوعاء، وهكذا... تُكرِّر هذه العملية ل $n$ مرَّة متتالية مع $m > n$ ( هذا النوع من التجارب يُعرَف باسم تجربة السحب بدون إعادة)، فإذا كنّا مهتمين بترتيب نتائج السحب، أي إنَّه (على سبيل المثال) يُنظر إلى تسلسل النتائج الآتية على أنَّها نتائج تختلف بعضها عن البعض الآخر:

$$\underbrace{1,2,3,4,5,6,7,...,k}_{n-\text{نتيجة}} \ \& \ \underbrace{2,1,3,4,5,6,7,...,k}_{n-\text{نتيجة}} \ \& \ \underbrace{3,2,1,4,5,6,7,...,k}_{n-\text{نتيجة}} \ \& ...$$

مع $2 \le k \le m$ ، فعندئذ سيكون لمجموعة نتائج هذه التجربة العشوائية العرض الآتي:

$$\Omega = \{ \omega : \omega = ( a_1 , a_2 ,..., a_n ) \ ; a_i \in \mathbb{N}_m \ \& \ i \in \mathbb{N}_n \ \& \ a_k \ne a_\ell \ for \ k \ne \ell \}$$

وفي هذه الحالة يكون لدينا:

$$|\Omega| = m(m-1)(m-2)\cdot ... \cdot (m-n+1)$$

وأمّا إذا كان الترتيب لعمليات السحب ليس له أهمية، أي إنَّه يُنظر إلى تسلسل النتائج الآتية على أنَّها تمثّل النتيجة نفسها:

$$\underbrace{1,2,3,4,5,6,7,...,k}_{n-\text{نتيجة}} \ \& \ \underbrace{2,1,3,4,5,6,7,...,k}_{n-\text{نتيجة}} \ \& \ \underbrace{3,2,1,4,5,6,7,...,k}_{n-\text{نتيجة}} \ \& ...$$

مع $2 \le k \le m$ ، فعندئذ يكون لمجموعة نتائج هذه التجربة العشوائية العرض الآتي:

$$\Omega = \{ \omega : \omega = [ a_1 , a_2 ,..., a_n ] \ ; a_i \in \mathbb{N}_m \ \& \ i \in \mathbb{N}_n \ \& \ a_k \ne a_\ell \ for \ k \ne \ell \}$$

ومن ثمَّ يكون عدد النتائج الممكنة لهذه التجربة يساوي:

$$|\Omega| = \binom{m}{n} = \frac{m!}{n!\,(m-n)!}$$

الجدول الآتي يوضِّح لنا المثالين الأخيرين من أجل الحالة الخاصّة $m = 3$ و $n = 2$ .

الجدول (٤ ,١)

| | | السحب | |
|---|---|---|---|
| | | مع الإعادة | بدون إعادة |
| الترتيب | له أهمية | (1,1) (1,2) (1,3)<br>(2,1) (2,2) (2,3)<br>(3,1) (3,2) (3,3) | (1,2) (1,3)<br>(2,1) (2,3)<br>(3,1) (3,2) |
| | عدد النتائج | $m^n = 3^2 = 9$ | $m(m-1)\cdots(m-n+1) = 3\times 2 = 6$ |
| | ليس له أهمية | [1,1] [2,2] [3,3]<br>[1,2] [1,3]<br>[2,3] | [1,2] [1,3]<br>[2,3] |
| | عدد النتائج | $\binom{m+n-1}{n} = \binom{4}{2} = 6$ | $\binom{m}{n} = \binom{3}{2} = 3$ |

**(٢, ١, ١, ٤) ملاحظة**

من الأمور المهمّة التي يجدر ذكرها في هذا الموضع هي أنَّ نتائج أمثلة سحب الكرات المُعطاة آنفاً تتوافق مع بعض التوزيعات لجسيمات دقيقة كائنة في حيِّز محدود، فعلى سبيل المثال نجد في الميكانيك الإحصائي مسألة توزيع $n$ جسيم (هذه الجسيمات قد تكون بروتونات، أو إلكترونات، أو فوتونات أو...) على $m$ حالة (هذه الحالات قد تكون على سبيل المثال مستويات الطاقة). إنَّ هذه الحالات يمكن تمثيلها بصناديق مرقَّمة بـ 1، 2، ... و$m$، فإذا افترضنا أنَّ الجسيمات متمايزة بعضها عن البعض الآخر، ومرقمة بـ 1، 2، ... و$n$، فعندئذ يمكن وصف توزيع الـ $n$ جسيم على الـ $m$ حالة بشكل تام من خلال عيِّنة مرتبة $(a_1, a_2, \ldots, a_n)$ (لمن يود الاطلاع على هذا النوع من العيِّنات يمكنه الرجوع إلى كتب نظرية المعاينة)، علماً أنَّ $a_i$ هو عدد الحالات التي مرَّ بها الجسيم ذي الرقم $i$. أمّا إذا كانت الجسيمات المأخوذة غير متمايزة، فعندئذ يمكن وصف توزيع الـ $n$ جسيم على الـ $m$ حالة بشكل تام من خلال عيِّنة غير مرتبة $[a_1, a_2, \ldots, a_n]$، علماً أنَّ $a_i$ هو عدد الحالات التي مرَّ بها الجسيم في المرَّة ذات الرقم $i$، وهكذا نلاحظ أنَّ:

العيِّنات المُرتَّبة **يقابلها** مسألة الجسيمات المتمايزة.

العيِّنات غير المُرتَّبة **يقابلها** مسألة الجسيمات غير المتمايزة.

وهذا يعني أنَّ حالة العيِّنات المرتبة لدى مسألة اختيار $n$ كرة من صندوق يحوي $m$ كرة هي ذات حالة تقسيم الجسيمات المتمايزة لدى مسألة توزيع $n$ جسيم على $m$ حالة، وبالمثل نجد أنَّ حالة العيِّنات غير المرتبة لدى مسألة اختيار $n$ كرة من صندوق يحوي $m$ كرة هي ذات حالة تقسيم الجسيمات غير المتمايزة لدى مسألة توزيع $n$ جسيم على $m$ حالة. هكذا نصل إلى النتيجة المهمّة الآتية (انظر [53]):

إنَّ السحب مع الإعادة يقابلها إمكانية وجود عدد كيفي من الجسيمات في حالة واحدة، وأمّا السحب بدون إعادة فيقابلها عدم إمكانية وجود أكثر من جسيم واحد في الحالة الواحدة.

إنَّ الفيزياء الإحصائي يُفيدنا بأنَّ الجسيمات المتمايزة التي لا تخضع لمبدأ باولي Pauli Principle (مبدأ الاستبعاد) ستكون مُحَقِّقَة إحصاءة ماكسويل – بولتسمان Maxwell–Boltzmann Statistic، وأمّا الجسيمات غير المتمايزة التي لا تخضع لمبدأ باولي ستكون مُحَقِّقَة إحصاءة بوزي – أينشتاين Bose – Einstein Statistic (ومن أمثال هذه الجسيمات الفوتونات والميزونات ذات النوع $\pi$). في حين أنّه إذا كانت الجسيمات غير متمايزة وتخضع لمبدأ باولي، فإنّها ستكون مُحَقِّقَة إحصاءة فيرمي – ديراك Fermi-Dirac Statistic (ومن أمثال هذه الجسيمات البروتونات والإلكترونات والنيترونات)، وأخيراً الجسيمات المتمايزة التي تخضع لمبدأ باولي فلا وجود لها في الفيزياء.

الجدول الآتي يوضِّح لنا أوجه التشابه بين مسألة سحب الكرات وتوزيعات الجسيمات الدقيقة.

الجدول (٢ , ٤)

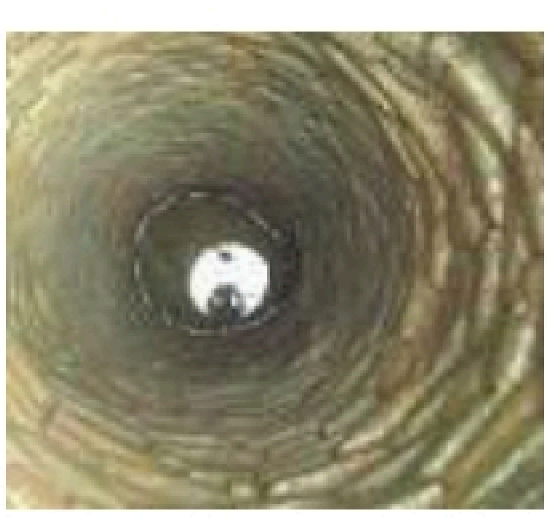
| | | | | | |
|---|---|---|---|---|---|
| | $\lvert\Omega\rvert$ لدى مسألة توزيع $n$ جسيم على $m$ حالة | | | | |
| | | الجسيمات | | | |
| | | غير متمايزة | متمايزة | | |
| السحب | مع الإعادة | $\binom{m+n-1}{n}$ إحصاءة بوزي – أينشتاين | $m^n$ إحصاءة ماكسويل – بولتسمان | بدون محدودية | التوزيع |
| السحب | بدون إعادة | $\binom{m}{n}$ إحصاءة فيرمي – ديراك | $m(m-1)\cdots\cdots(m-n+1)$ | مع المحدودية | التوزيع |
| | | غير مرتبة | مرتبة | | |
| | | لعيِّنة | | | |
| | $\lvert\Omega\rvert$ لدى مسألة سحب $n$ كرة من وعاء يحوي $m$ كرة | | | | |

**(٤ , ١ , ١ , ٣) أمثلة على تجارب ذات عدد غير منته من النتائج:**

١ – يقذف شخص قطعة نقود، فإذا حصل على صورة ينهي التجربة، وأمّا إذا حصل على شعار فإنَّه سيقذف قطعة النقود مرَّةً أخرى، فإن حصل على صورة ينهي التجربة وإلاّ سيعيد قذف قطعة النقود من جديد. عندئذ ستكون لمجموعة نتائج هذه التجربة العشوائية العرض الآتي:

$$\Omega=\{H,(T,H),(T,T,H),(T,T,T,H),\ldots\}$$

فنجد أنَّ $\Omega$ مجموعة غير منتهية ولكنَّها قابلة للعدّ.



٢– يقف شخص على حافة بئر دائري الشكل بنصف قطر $R$، ويلقي بحصى إلى قاع البئر، فعندئذ نجد أن كل نقطة $(x,y)$ من سطح ماء البئر ستكون نتيجةً محتملةً للإصابة، ومن ثمَّ ستكون لمجموعة نتائج هذه التجربة العشوائية العرض الآتي:

$$\Omega=\left\{(x,y)\in\mathbb{R}^2\ ;x^2+y^2\le R^2\right\}$$

وهنا نلاحظ أنَّ $\Omega$ مجموعة غير منتهية، وأكثر من ذلك غير قابلة للعدّ أيضاً.

**والخلاصة:**

يتضِّح لنا مّما سبق أنَّه لا توجد قاعدة محدَّدة لتعيين مجموعة نتائج تجربة عشوائية، وإنَّما ذلك يتعلَّق بطبيعة التجربة والشروط المفروضة عليها. لذلك يجب إمعان النظر في نصوص مسائل الاحتمالات المقدَّمة لكي نصل إلى نتائج صحيحة، وهذا التحذير واجب في الاحتمالات أكثر منه في بقية الاختصاصات، وذلك لأنَّ الكثير من مسائل الاحتمالات تُقدَّم على شكل جُمل أو نصوص وليس في شكل علاقات رياضياتية.

**(٤ , ١ , ٢) جبور الحوادث Algebras of Events**

نعلم أنَّ كل تجربة عشوائية تتبعها مجموعة معينة من النتائج الممكنة، وهذه النتائج تُدعى **حوادث ابتدائية** Elementary Events (أو حوادث ذرية **Atomic Events**) ولهذا السبب يُطلق على مجموعة **كل** نتائج التجربة العشوائية $\Omega$ اسم **فضاء الحوادث الابتدائية** Space of Elementary Events أيضاً.

الآن، ومن أجل تجربة عشوائية ما يجب علينا البتّ في النتيجة المطلوبة إن كانت قد تحقَّقت أم لا، وذلك من أجل مقولة مُحدَّدة متعلِّقة بهذه التجربة. هذا يكافئ القول: إنَّ المجرِّب لا يهتمُّ بأية نتيجة للتجربة سيأخذ، وإنَّما الذي يهمُّه هو إن كانت هذه النتيجة ستنتمي إلى مجموعة جزئية أو أكثر من مجموعة كل نتائج التجربة. إنَّ أمثال هذه المجموعات الجزئية تعرف باسم **حوادث** Events، وقد قدَّم الرياضياتي والعشوائي الروسي **شيرياييف** (Albert Nikolaevich Shiryaev (Born 1934 (**عالم رياضيات وعشوائيات**) مفهوم الحادث على النحو السابق ذكره، وهذا المفهوم مطابق للتعريف الذي قدَّمه الإحصائي والرياضياتي الفرنسي **لوفي** Michel Loève (1907-1979) أيضاً (**انظر** [53] و[39]).

فيما يلي سنقدِّم بعض المفاهيم المتعلِّقة بالحوادث المركَّبة حيث نعلم أنَّه في علـم المنطـق يمكننا تشكيل مقولات مركَّبة بوساطة العمليّات المنطقيّة (**و**)، (**أو**) و(**كلا**) التي يقابلها في نظرية المجموعات **التقاطع** ($\cap$)، **الاتحاد** ($\cup$) و **الفرق** ($\setminus$) على الترتيب، وبما أنَّ الحوادث الناتجة عن التجارب العشوائية تُمثَّل في الرياضيات بمجموعات فإنَّ ذلك يبرِّر ظهور حوادث مركَّبة وفقاً للعمليّات المنطقيّة على المجموعات.

### (١, ٢, ١, ٤) تعاريف

لتكن لدينا تجربة عشوائية مجموع نتائجها $\Omega$، ولنفترض أنَّ $A$ و$B$ حادثين متعلِّقين بهذه التجربة (**أي إنَّ** $A\ ,B \subseteq \Omega$)، فعندئذ:

١- حادث تحقُّق $A$ **أو** $B$ (**الذي يُرمز له بـ** $A \cup B$) هو ذلك الحادث الذي يُعبِّر عن انتماء نتيجة التجربة $\omega$ إلى $A$ أو $B$ أو إلى كليهما.

٢- حادث تحـقُّق $A$ **و** $B$ معاً (**الذي يُرمز له بـ** $A \cap B$) هو ذلك الحادث الذي يُعبِّر عن انتماء نتيجة التجربة $\omega$ إلى كلٍّ من $A$ و$B$ في آنٍ واحد.

٣- حادث تحقُّق $A$ ولكن دون $B$ (**والذي يُرمز له بـ** $A \setminus B$)، هو ذلك الحادث الذي يُعبِّر عن انتماء نتيجة التجربة $\omega$ إلى الحادث $A$ ولكن دون أن تنتمي إلى الحادث $B$.

### (٢, ٢, ١, ٤) ملاحظات

١- في الحالة الخاصّة إذا كان حادث $A \subseteq \Omega$ يحوي نتيجة واحدة فقط من نتائج التجربة العشوائية، كأن يكون $A = \{\omega\}$، فعندئذٍ يُقال عن هذا الحادث إنَّه **حادثٌ بسيط** Simple Event.

٢- في الحقيقة يمكننا أن نشتق عدداً كبيراً من الحوادث اعتماداً على العمليّات الثلاث السّابقة، ومن أهمِّها الحوادث الآتية:

أ) **متمِّم حادث** **Complement of an Event**: ليكن $A \subseteq \Omega$ حادثاً مُعطى، فعندئذ **الحادث المتمِّم** لـ $A$ (**وسنُرمز له بـ** $\overline{A}$) يُعرَّف على أنَّه الحادث الذي يتحقَّق في حال عدم تحقُّق الحادث $A$، وهذا يعني أنَّه إذا كانت نتيجة التجربة $\omega$ لا تنتمي إلى $A$ فإنَّها ستنتمي إلى مُتمِّم $A$ حتماً. أي إنَّ $\overline{A} = \Omega \setminus A$.

ب) **الحادث الأكيد** **Certain Event**: إنَّ هذا الحادث موافق للحقيقة القائلة: إنَّه من أجل أي حادث $A \subseteq \Omega$ ستكون أية نتيجة $\omega$ من نتائج التجربة منتمية إلى الحادث $A \cup \overline{A}$ بكل تأكيد، ولذلك يُدعى أمثال هذا الحادث بـ **الحادث الأكيد**، وبما أنَّ $A \cup \overline{A} = \Omega$ فإنَّه يُرمز للحادث الأكيد بـ $\Omega$ أيضاً.

ج) **الحادث المستحيل** **Impossible Event**: إنَّ هذا الحادث متأتٍ عن حقيقة أنَّه من أجل أي حادث $A \subseteq \Omega$، فإنَّ نتيجة ما $\omega$ من نتائج التجربة لا يمكن أن تنتمي إلى الحادث $A \cap \overline{A}$ ولا بأي شكل من الأشكال، أي إنَّه من المستحيل أن تنتمي $\omega$ إلى الحادث $A \cap \overline{A}$، ولذلك يُدعى أمثال هذا الحادث بـ **الحادث المستحيل**، وبما أنَّ $A \cap \overline{A} = \varnothing$ فإنَّه يُرمز للحادث المستحيل بـ $\varnothing$ أيضاً.

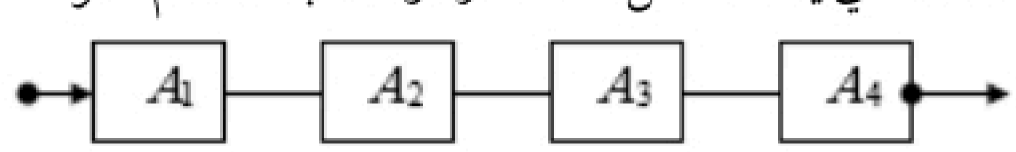
د ) الحوادث المتنافية **Mutually Exclusive Events**: يصادفنا في كثير من الحالات أنَّ انتهاء نتيجة التجربة إلى حادثين $A, B \subseteq \Omega$ بآن واحد أمر غير ممكن (مستحيل)، أي إنَّ $A \cap B = \varnothing$ ، فعندئذٍ يُقال إنَّ الحادثين $A$ و$B$ **متنافيان**، أو غير متلائمين، أو بعضه مانع لبعضه الآخر. أمّا إذا كان لدينا متتالية حوادث $(A_i)_{i \in I}$ من $\Omega$ مع $I$ مجموعة قابلة للعدّ، فعندئذ إذا كان من أجل كل القيم الممكنة ل $i$ و$j$ لدينا $A_i \underset{i \neq j}{\cap} A_j = \varnothing$ فعندئذٍ يُقال عن متتالية الحوادث $(A_i)_{i \in I}$ إنَّها **متنافية مثنى مثنى** Pairwise Mutually Exclusive.

٣- إذا كان لدينا $(A_i)_{i \in I}$ متتالية حوادث متنافية مثنى مثنى، فإنَّه يُكْتب غالباً (وعلى سبيل التوضيح) $\sum_{i \in I} A_i$ عوضاً عن $\bigcup_{i \in I} A_i$ ، وفي الحالة الخاصَّة يكتب $A + B$ عوضاً عن $A \cup B$ من أجل $A$ و$B$ حادثين متنافيين.

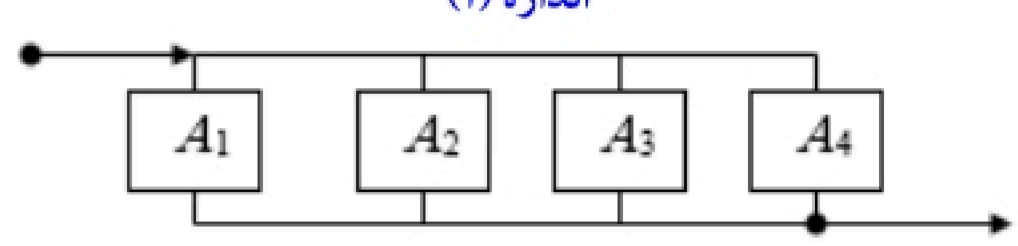
٤- بناءً على ما سبق نلاحظ أنَّ أي حادث $A = \{ \omega_i \; ; \; i \in I \}$ من $\Omega$ يمكن النظر إليه كحادث مُركَّب من حوادث بسيطة وذلك لأنَّه يُمكننا أن نكتب $A = \sum_{i \in I} \{ \omega_i \}$ أيضاً.

**(٣, ٢, ١, ٤) أمثلة:**

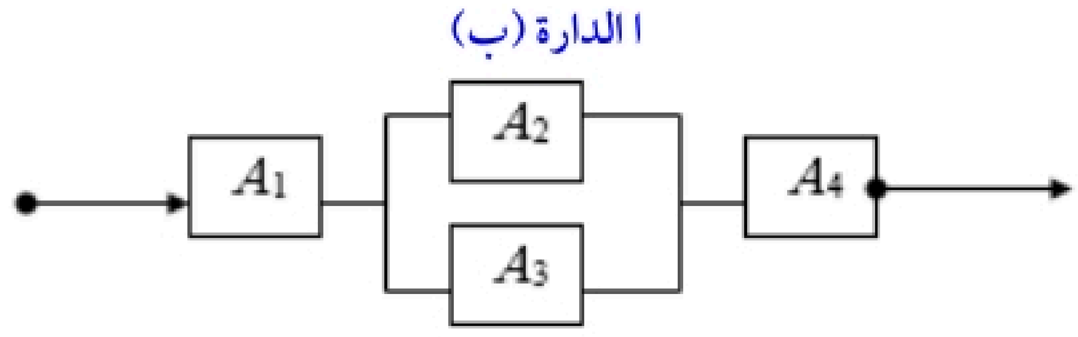
لدينا دارة كهربائية مكوَّنة من أربع قطع تؤدي النوع نفسه من العمل (دارات مفاتيح)، فإذا رمزنا بـ $A_i$ مع $i \in \mathbb{N}_4$ للحادث الذي يعبِّر عن عمل القطعة ذات الرقم $i$ ، وبفرض أنَّ الإشارة ستمرُّ من مسار واحد فقط، وأنَّنا أدخلنا إشارة في مدخل الدارة المفترضة، فعندئذ لنعبِّر عن الحادث الذي يصف عمل هذه الدارة وذلك باستخدام الحوادث $A_i$ .



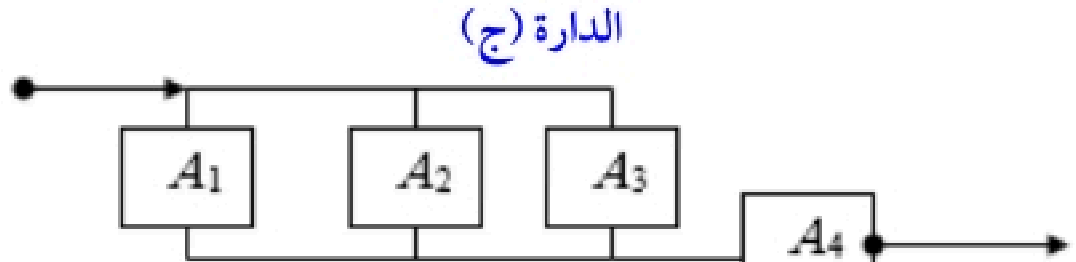

الدارة (أ)



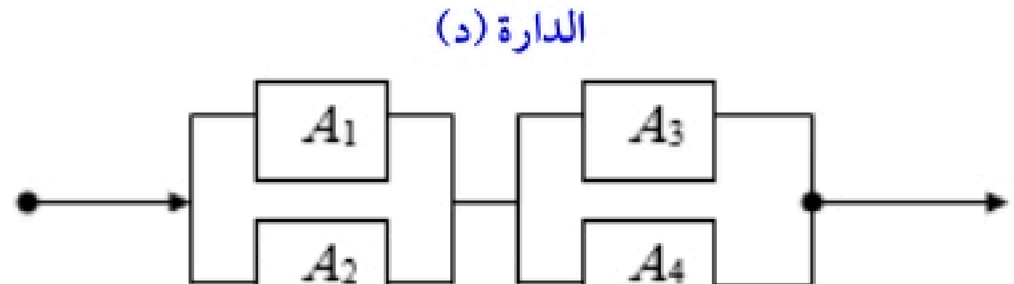

ا الدارة (ب)




الدارة (ج)




الدارة (د)




الدارة (هـ)

**الحلول**: من أجل الدارة:

أ) نلاحظ أنَّ هذه الدارة تؤدي عملها إذا قامت جميع القطع بأداء عملها، ومن ثمَّ بفرض أنَّ $A$ هو الحادث الذي يعبِّر عن عمل هذه الدارة فإنَّه يمكننا كتابة الحادث $A$ باستخدام الحوادث $A_i$ وفقاً للعلاقة الآتية:

$$A = A_1 \cap A_2 \cap A_3 \cap A_4$$

ب) نلاحظ أنَّ هذه الدارة تؤدي عملها إذا قامت إحدى هذه القطع بأداء عملها، ومن ثمَّ بفرض أنَّ $B$ هو الحادث الذي يعبِّر عن عمل هذه الدارة فإنَّه يمكننا كتابة الحادث $B$ باستخدام الحوادث $A_i$ وفقاً للعلاقة الآتية:

$$B = A_1 \cup A_2 \cup A_3 \cup A_4$$

ج) نلاحظ أنَّ هذه الدارة تؤدي عملها إذا قامت القطع $A_1$ و $A_4$ وإحدى القطعتين $A_2$ أو $A_3$ بأداء عملها، ومن ثمَّ فإنَّه بفرض أنَّ $C$ هو الحادث الذي يعبِّر عن عمل هذه الدارة فإنَّه يمكننا كتابة الحادث $C$ باستخدام الحوادث $A_i$ وفقاً للعلاقة الآتية:

$$C = A_1 \cap (A_2 \cup A_3) \cap A_4 = (A_1 \cap A_2 \cap A_4) \cup (A_1 \cap A_2 \cap A_4)$$

د) نلاحظ أنَّ هذه الدارة تؤدي عملها إذا قامت القطعة $A_4$ وإحدى القطعتين $A_1$ أو $A_2$ أو $A_3$ بأداء عملها، ومن ثمَّ فإنَّه بفرض أنَّ $D$ هو الحادث الذي يعبِّر عن عمل هذه الدارة فإنَّه يمكننا كتابة الحادث $D$ باستخدام الحوادث $A_i$ وفقاً للعلاقة الآتية:

$$D = (A_1 \cup A_2 \cup A_3) \cap A_4 = (A_1 \cap A_4) \cup (A_2 \cap A_4) \cup (A_3 \cap A_4)$$

هـ) نلاحظ أنَّ هذه الدارة تؤدي عملها إذا قامت إحدى القطعتين $A_1$ أو $A_2$ مع إحدى القطعتين $A_3$ أو $A_4$ بأداء عملها، ومن ثمَّ فإنَّه بفرض أنَّ $E$ هو الحادث الذي يعبِّر عن عمل هذه الدارة فإنَّه يمكننا كتابة الحادث $E$ باستخدام الحوادث $A_i$ وفقاً للعلاقة الآتية:

$$E = (A_1 \cup A_2) \cap (A_3 \cup A_4) = (A_1 \cap A_3) \cup (A_1 \cap A_4) \cup (A_2 \cap A_3) \cup (A_2 \cap A_4)$$

لقد لاحظنا سابقاً أنَّ كل مقولة متعلقَّة بتجربة عشوائية هناك حادث يمثِّلها، والعكس صحيح أيضاً، فمن أجل أي حادث يمكن صياغة مقولة تعبِّر عنه، فعلى سبيل المثال لو أخذنا تجربة قذف قطعة نقود معدنية لمرّتين متتاليتين، وكنّا مهتمين بحادث الحصول على صورة واحدة فقط، فعندئذ يُعبَّر عن ذلك رياضياتياً بالمجموعة $A = \{HT, TH\}$. الآن، وبشكل معكوس لو أُعطينا المجموعة $B = \{HH\}$، فعندئذ نعبِّر عن ذلك بالقول: إنَّه قد تحقَّق لدينا حاث الحصول على صورتين تماماً في هذه التجربة العشوائية. لكن نعلم أنَّ المقولات تخضع لجبر المنطق (جبر بول **Pool algebra**) ولذلك فإنَّ الحوادث المقابلة لهذه المقولات يجب أن تكون منسجمة مع العلاقات المنطقية أيضاً، ومن ثمَّ مع شروط جبر المنطق. أي يجب أن تكون خصائص صف الحوادث لتجربة عشوائية ما مُكافئةً لخصائص جبر المنطق، ومن ثمَّ العلاقات بين الحوادث سوف تصاغ وتقبل على شكل مسلَّمات موافقة لجبر المنطق أيضاً. إنَّ هذا الحوار المنطقي دفع الرياضياتي غليفينكو إلى تقديم ما يُعرف باسم جبر الحوادث الذي عدَّه ركيزة أساس في الحساب الاحتمالي عندما قدَّم مقالته الموسومة بـ **مدخل إلى الحساب الاحتمالي** عام 1939.

**(٤ ,١ ,٢ ,٤) ملاحظة**

من أجل $\Omega$ مجموعة نتائج تجربة عشوائية ما سنرمز بـ $2^{\Omega}$ لأسرة كل المجموعات الجزئية من $\Omega$.

و الآن سنقدِّم تعريف ما يُسمَّى بـ **الجبر** والـ **$\sigma$ – جبر** فوق $\Omega$.

**(٤ ,١ ,٢ ,٥) تعريف (الجبر فوق Algebra on Ω)**

لتكن Ω مجموعة نتائج تجربة عشوائية، ولنأخذ $\mathscr{A}$ صف مجموعات جزئي من أسرة كل المجموعات الجزئية لـ Ω (أي إنَّ $\mathscr{A} \subseteq 2^{\Omega}$)، ولنفترض البنود الآتية:

$\mathbf{a}_1$ – المجموعة Ω عنصر من $\mathscr{A}$ (أي إنَّ $\Omega \in \mathscr{A}$).

$\mathbf{a}_2$ – من أجل أي $A$ و$B$ من $\mathscr{A}$ فإنَّ $A \cup B$ هو عنصر من $\mathscr{A}$ أيضاً (أي إنَّ الصف $\mathscr{A}$ مغلق بالنسبة إلى عملية الاتحاد المنتهي).

$\mathbf{a}_3$ – من أجل أي $A$ من $\mathscr{A}$ فإنَّ $\overline{A}$ هو عنصر من $\mathscr{A}$ أيضاً (أي إنَّ الصف $\mathscr{A}$ مغلق بالنسبة إلى عملية التتميم).

$\mathbf{a}_\infty$ – من أجل أية متتالية $(A_n)_{n\in\mathbb{N}}$ عناصرها من $\mathscr{A}$ فإنَّ $\bigcup_{n=1}^{\infty} A_n$ هو عنصر من $\mathscr{A}$ أيضاً (أي إن الصف $\mathscr{A}$ مغلق بالنسبة إلى عملية الاتحاد غير المنتهي العدود).

عندئذ يُقال إنَّ:

أ ) الصف $\mathscr{A}$ هو جبر Algebra فوق Ω إذا حَقَّق هذا الصف البنود $(\mathbf{a}_1)$ و $(\mathbf{a}_2)$ و $(\mathbf{a}_3)$.

ب) الصف $\mathscr{A}$ هو σ – **جبر** σ-Algebra فوق Ω إذا حَقَّق هذا الصف البنود $(\mathbf{a}_1)$ و $(\mathbf{a}_\infty)$ و $(\mathbf{a}_3)$.

**(٤ ,١ ,٢ ,٦) ملاحظة**

إذا ذُكِر أنَّ الصف $\mathscr{A}$ غير خال فإنَّ ذلك يعني أنَّه يوجد على الأقل عنصر $A \in \mathscr{A}$، ولذلك فإذا كان $\mathscr{A}$ جبراً (أو σ – **جبر**) فإنَّه بسبب تَحَقُّق البند $(\mathbf{a}_3)$ سيكون لدينا $\overline{A} \in \mathscr{A}$، وكذلك بحسب البند $(\mathbf{a}_2)$ (أو $(\mathbf{a}_\infty)$) سينتج لدينا أنَّ $A \cup \overline{A} = \Omega$ هو عنصر من $\mathscr{A}$ أيضاً، ولذلك يمكن حذف البند الأول $(\mathbf{a}_1)$ في هذه الحالة (أي إذا أُخِذَ الجبر $\mathscr{A}$ كصف غير خال).

**(٤ ,١ ,٢ ,٧) مثال**

لنأخذ تجربة قذف قطعة نقود معدنية لمرّة واحدة، فعندئذ نعلم أنَّ مجموعة نتائج هذه التجربة العشوائية هي $\Omega = \{H, T\}$، ومن ثمَّ سيكون لأسرة كل المجموعات الجزئية من Ω العرض الآتي:

$$2^{\Omega} = \{\varnothing, \{H\}, \{T\}, \{H,T\} = \Omega\}$$

فلو أخذنا:

$$\mathscr{A}_1 = \{\varnothing, \{H\}, \{T\}, \Omega\} = 2^{\Omega} \qquad \& \qquad \mathscr{A}_2 = \{\varnothing, \{T\}, \Omega\}$$

$$\mathscr{A}_3 = \{\varnothing, \{H\}, \Omega\} \qquad \& \qquad \mathscr{A}_4 = \{\varnothing, \Omega\}$$

فنجد وبكل وضوح أنَّ كلًّا من $\mathscr{A}_1$ و $\mathscr{A}_4$ هو جبر فوق Ω، وأمّا $\mathscr{A}_2$ و $\mathscr{A}_3$ فإنَّهما لا يحقّقان شروط الجبر فوق Ω (شرط الإغلاق بالنسبة إلى عملية التتميم غير محقّق). كذلك نجد أنَّ الصف $\mathscr{A}_1$ هو σ – **جبر** على Ω، وذلك لأنَّه مغلق بالنسبة إلى عملية التتميم وضوحاً، وكذلك مغلق بالنسبة إلى عمليّة الاتحاد غير المنتهي العدود لأنَّه:

- لو أخذنا $A_n = \varnothing$ من أجل كل $n \in \mathbb{N}$، فعندئذ سيكون لدينا $\bigcup_{n=1}^{\infty} A_n = \varnothing$.
- لو أخذنا $A_n = \Omega$ من أجل كل $n \in \mathbb{N}$، فعندئذ سيكون لدينا $\bigcup_{n=1}^{\infty} A_n = \Omega$.
- لو أخذنا $A_n = \varnothing$ من أجل بعض القيم لـ $n \in \mathbb{N}$ سواء كان عددها منتهيًا أم غير منته، وباقي عناصر المتتالية $(A_n)_{n\in\mathbb{N}}$ هي $\{H\}$، فعندئذ سيكون لدينا $\bigcup_{n=1}^{\infty} A_n = \{H\}$.
- لو أخذنا $A_n = \varnothing$ من أجل بعض القيم لـ $n \in \mathbb{N}$ سواء كان عددها منتهيًا أم غير منته، وباقي عناصر المتتالية $(A_n)_{n\in\mathbb{N}}$ هي $\{T\}$، فعندئذ سيكون لدينا $\bigcup_{n=1}^{\infty} A_n = \{T\}$.

- لو أخذنا $A_n = \varnothing$ من أجل بعض القيم لـ $n \in \mathbb{N}$ سواء كان عددها منتهيًا أم غير منته، وباقي عناصر المتتالية $(A_n)_{n\in\mathbb{N}}$ مكونة من $\{H\}$ و $\{T\}$، فعندئذ سيكون لدينا $\bigcup_{n=1}^{\infty} A_n = \Omega$ .
- لو أخذنا $A_n = \Omega$ من أجل بعض القيم لـ $n \in \mathbb{N}$ سواء كان عددها منتهيًا أم غير منته، وباقي عناصر المتتالية $(A_n)_{n\in\mathbb{N}}$ هي $\{H\}$، فعندئذ سيكون لدينا $\bigcup_{n=1}^{\infty} A_n = \Omega$ .
- لو أخذنا $A_n = \Omega$ من أجل بعض القيم لـ $n \in \mathbb{N}$ سواء كان عددها منتهيًا أم غير منته، وباقي عناصر المتتالية $(A_n)_{n\in\mathbb{N}}$ هي $\{H\}$، فعندئذ سيكون لدينا $\bigcup_{n=1}^{\infty} A_n = \Omega$ .
- لو أخذنا $A_n = \Omega$ من أجل بعض القيم لـ $n \in \mathbb{N}$ سواء كان عددها منتهيًا أم غير منته، وباقي عناصر المتتالية $(A_n)_{n\in\mathbb{N}}$ مكونة من $\{H\}$ و $\{T\}$، فعندئذ سيكون لدينا $\bigcup_{n=1}^{\infty} A_n = \Omega$ .
- لو أخذنا $A_n = \{H\}$ من أجل كل $n \in \mathbb{N}$، فعندئذ سيكون لدينا $\bigcup_{n=1}^{\infty} A_n = \{H\}$ .
- لو أخذنا $A_n = \{T\}$ من أجل كل $n \in \mathbb{N}$، فعندئذ سيكون لدينا $\bigcup_{n=1}^{\infty} A_n = \{T\}$ .
- لو أخذنا $A_n = \{H\}$ من أجل بعض القيم لـ $n \in \mathbb{N}$ سواء كان عددها منتهيًا أم غير منته، وباقي عناصر المتتالية $(A_n)_{n\in\mathbb{N}}$ هي $\{T\}$، فعندئذ سيكون لدينا $\bigcup_{n=1}^{\infty} A_n = \Omega$ .
- لو أخذنا $A_n = \{T\}$ من أجل بعض القيم لـ $n \in \mathbb{N}$ سواء كان عددها منتهيًا أم غير منته، وباقي عناصر المتتالية $(A_n)_{n\in\mathbb{N}}$ هي $\{H\}$، فعندئذ سيكون لدينا $\bigcup_{n=1}^{\infty} A_n = \Omega$ .

حيث نعلم أنَّ $\varnothing$، $\Omega$، $\{H\}$، و $\{T\}$ جميعها عناصر من $\mathscr{A}$ .

أمّا من أجل $\mathscr{A}_4$ فإنَّنا نجد بمناقشة مماثلة (وأكثر بساطة) أنَّه $\sigma$ - **جبر** على $\Omega$ أيضاً، وأخيراً نلاحظ أنَّ $\mathscr{A}_2$ و $\mathscr{A}_3$ لا يحقّقان شروط الـ $\sigma$ - **جبر** على $\Omega$ للسبب الذي ذكرناه سابقا (بخصوص الإغلاق بالنسبة إلى عملية التتميم).

**(٤ , ١ , ٢ , ٨) ملاحظة**

من أجل $\Omega$ مجموعة نتائج تجربة عشوائية، فإنَّ عناصر الصف $\mathscr{A}$ (جبراً كان أم $\sigma$ - **جبر** على $\Omega$) هي حوادث متعلّقة بالتجربة العشوائية التي قيد الدرس.

**(٤ , ١ , ٢ , ٩) نتائج**

١ - إنَّ كل جبر فوق $\Omega$ يكون مغلقاً بالنسبة إلى عملية التقاطع المنتهي، أي إنَّه من أجل كل $A_1$، $A_2$، ... و $A_n$ من $\mathscr{A}$ فإنَّ $\bigcap_{i=1}^{n} A_i$ هو حادث من $\mathscr{A}$ أيضاً.

**الإثبات**: لإثبات ذلك يكفي أن نبين أنَّه من أجل أي حادثين $A, B \in \mathscr{A}$ يكون $A \cap B \in \mathscr{A}$ أيضاً، وذلك لأنَّه إذا أخذنا $A_1 = A$ و $A_2 = B$ وأثبتنا أنَّ $A \cap B \in \mathscr{A}$ فإنَّنا نضع من جديد $A := A_1 \cap A_2$ ونأخذ $B =: A_3$ فيكون لدينا من جديد $A \cap B \in \mathscr{A}$ أيضاً، وهكذا دواليك.

الآن لإثبات أنَّ $A \cap B \in \mathscr{A}$ سننطلق من أنَّ $A, B \in \mathscr{A}$ حادثين كيفيين، فيكون بحسب الشرط $\mathbf{a}_3$ لدينا $\overline{A}$ و $\overline{B}$ من $\mathscr{A}$، وبحسب الشرط $\mathbf{a}_2$ ينتج أنَّ $\overline{A} \cup \overline{B}$ من $\mathscr{A}$ ، ومرَّة أخرى بحسب الشرط $\mathbf{a}_3$ سيكون $\overline{\overline{A} \cup \overline{B}}$ من $\mathscr{A}$ أيضاً، ومن ثمّ نحصل على المطلوب بتطبيق قانون ديمورغان (انظر الملحق **A**) حيث لدينا $A \cap B = \overline{\overline{A} \cup \overline{B}}$ .

٢- إنَّ كل $\sigma$- **جبر** على $\Omega$ يكون مغلقاً بالنسبة إلى عملية التقاطع غير المنتهي العدود، أي إنَّه بفرض أنَّ $\mathscr{A}$ هو $\sigma$- **جبر** على $\Omega$، فعندئذ من أي متتالية حوادث $(A_n)_{n\in\mathbb{N}}$ عناصرها من $\mathscr{A}$ فإنَّ $\bigcap_{n=1}^{\infty} A_n$ هو عنصر من $\mathscr{A}$ أيضاً.

**الإثبات:** لتكن $(A_n)_{n\in\mathbb{N}}$ متتالية حوادث من $\mathscr{A}$، فعندئذ باستخدام قانون ديمورغان وبحسب تعريف الـ $\sigma$- **جبر** يكون لدينا $\bigcap_{i=1}^{\infty} A_i = \overline{\bigcup_{i=1}^{\infty} \overline{A_i}}$ وهو عنصر من $\mathscr{A}$.

٣- إذا كان $\mathscr{A}$ جبراً على $\Omega$، فإنَّ كلا الحادثين المستحيل $\varnothing$ والأكيد $\Omega$ ينتميان إلى $\mathscr{A}$ أيضاً.

**الإثبات:** بما أنَّ $\mathscr{A}$ صف غير خال فإنَّه يوجد على الأقل حادث $A \in \mathscr{A}$، ومنه يكون لدينا $\overline{A} \in \mathscr{A}$ أيضاً، ومن ثمَّ ينتج لدينا بحسب البند الأول من تعريف الجبر أنَّ $A \cup \overline{A} = \Omega$ هو عنصر من $\mathscr{A}$، وكذلك بحسب الفقرة السابقة /١/ يكون لدينا $A \cap \overline{A} = \varnothing$ من $\mathscr{A}$ أيضاً. لاحظ هنا أنَّه من الممكن أن يكون $A = \Omega$، ومن ثمَّ ينتج مباشرة أنَّ $\varnothing$ و $\Omega$ هي عناصر من $\mathscr{A}$ أيضاً.

٤- بشكل مماثل لما سبق يمكننا إثبات أنَّه إذا كان $\mathscr{A}$ يشكِّل $\sigma$- **جبر** فوق $\Omega$ فإنَّ الحادثين $\varnothing$ و $\Omega$ ينتميان إلى $\mathscr{A}$ أيضاً.

٥- إنَّ كلاً من الجبر والـ $\sigma$- **جبر** فوق $\Omega$ يكون مغلقاً بالنسبة إلى عمليّة الفرق بين الحوادث.

**الإثبات:** ليكن $A$ و $B$ حادثين من $\mathscr{A}$، فعندئذ بحسب تعريف الجبر (**أو الـ $\sigma$- جبر**) والفقرات السابقة سيكون لدينا $A \cap \overline{B} = A \setminus B$ وهو عنصر من $\mathscr{A}$.

٦- لتكن $\Omega$ مجموعة نتائج تجربة عشوائية، فعندئذ كل $\sigma$- **جبر** فوق $\Omega$ هو جبر فوق $\Omega$ أيضاً.

**الإثبات:** يكفي من أجل ذلك التحقّق من البند الأول في تعريف الجبر لأنَّ البند الثاني مشترك بينهما، فلو افترضنا أنَّ $\mathscr{A}$ هو $\sigma$- **جبر** فوق $\Omega$، ولتكن $(A_n)_{n\in\mathbb{N}}$ متتالية ما من الحوادث في $\mathscr{A}$ وبحيث يكون فيها حادثان $A_i \neq \varnothing$ و $A_j \neq \varnothing$ مع $i \neq j$ قيمتين صحيحتين مثبّتتين، وأمّا باقي الحوادث فهي $\varnothing$، فعندئذ سيكون $\bigcup_{n=1}^{\infty} A_n = A_i \cup A_j$ من $\mathscr{A}$، وهذا يعني أنَّ $\mathscr{A}$ هو جبر فوق $\Omega$.

٧- لقد لاحظنا في مثال سابق أن بعض الجبور قد تكون $\sigma$- **جبر** أيضاً (انظر المثال (٤،١،٢،٧))، ولكن ليس كل جبر $\mathscr{A}$ فوق $\Omega$ هو $\sigma$- **جبر** فوق $\Omega$.

**الإثبات:** من أجل ذلك لنأخذ مجموعة نتائج التجربة العشوائية $\Omega$ غير منتهية ولكن قابلة للعدّ على، وليكن $\mathscr{A} \subseteq 2^{\Omega}$ مُعطى كما يلي:

$$\mathscr{A} := \left\{ A \in 2^{\Omega} \;;\; A \text{ finite or } \overline{A} \text{ finite} \right\} \qquad [4,1]$$

فعندئذ نجد أنَّ هذا الصف $\mathscr{A}$ هو جبر فوق $\Omega$، ويسمّى **جبر المتمّمات المنتهية** Algebra of Finite Complements، حيث لدينا:

١- إذا كان $A \in \mathscr{A}$ فإنَّ $\overline{A} \in \mathscr{A}$ أيضاً، وذلك للأسباب الآتية:

أ- إذا كان الحادث $A$ منتهياً فإنَّ الحادث $\overline{A}$ يملك متمماً منتهياً لأنَّ $A = \overline{\overline{A}}$.

ب- إذا كان الحادث $A$ يملك متمماً منتهياً فإنَّ $\overline{A}$ يكون منتهياً.

٢- ليكن $A$ و $B$ حادثين من $\mathscr{A}$ فعندئذ سيكون $A \cup B \in \mathscr{A}$ أيضاً، وذلك للأسباب الآتية:

أ) إذا كان الحادثان $A$ و $B$ منتهيين فإنَّ الحادث $A \cup B$ سيكون منتهياً.

ب- إذا كان أحد الحادثين $A$ أو $B$ منتهياً والآخر يملك متمماً منتهياً، ولنفترض على سبيل المثال أنَّ الحادث $A$ منتهياً و $B$ يملك متمماً منتهياً، فإنَّ الحادث $\overline{A \cup B} = \overline{A} \cap \overline{B} \subseteq \overline{B}$ سيكون منتهياً أيضاً، ومن ثمَّ فإنَّ $A \cup B$ يملك متمماً منتهياً.

ج- إذا كان كلٌّ من الحادثين $A$ أو $B$ يملك متمماً منتهياً، فعندئذ سيكون الحادث $\overline{A \cup B} = \bar{A} \cap \bar{B} \subseteq \bar{A}$ منتهياً أيضاً، ومن ثمّ الحادث $A \cup B$ يملك متمماً منتهياً.

لكنَّ هذا الجبر $\mathscr{A}$ **ليس** $\sigma$- **جبر** فوق $\Omega$ وذلك لأنَّه لو أخذنا (على سبيل المثال) $\Omega = \mathbb{Z}$، وأخذنا متتالية الحوادث $(\{n\})_{n\in\mathbb{N}}$ التي عناصرها من $\mathscr{A}$، فإنَّنا نجد الآتي:

$$\bigcup_{n=1}^{\infty}\{n\} = \mathbb{N} \notin \mathscr{A} \qquad \& \qquad \overline{\bigcup_{n=1}^{\infty}\{n\}} = \mathbb{Z}\setminus\mathbb{N} \notin \mathscr{A}$$

وبهذا يتمُّ الإثبات.

٨- في الحالة الخاصّة عندما يكون فضاء الحوادث الابتدائية $\Omega$ منتهياً، فإنَّ كل جبر فوق $\Omega$ هو $\sigma$- **جبر** فوق $\Omega$ أيضاً.

**الإثبات**: لنفترض أنَّ فضاء الحوادث الابتدائية $\Omega$ منتهياً، وليكن $\mathscr{A} \subseteq 2^{\Omega}$ جبراً فوق $\Omega$، فعندئذ لإثبات أنَّ $\mathscr{A}$ هو $\sigma$- **جبر** فوق $\Omega$ يكفي إثبات الإغلاق بالنسبة إلى عملية الاتحاد غير المنتهي العدود.

بما أنَّ فضاء الحوادث الابتدائية $\Omega$ منتهياً فإنَّ عدد الحوادث في $\mathscr{A}$ سيكون منتهياً حتماً، ومنه بأخذ الحوادث $A_1$، $A_2$، ... و $A_n$ من $\mathscr{A}$ مع $n$ عدد طبيعي محدود بحيث يكون $|\mathscr{A}| \leq n$، فإنَّه يمكننا أن نشكِّل متتالية غير منتهية $(B_i)_{i\in\mathbb{N}}$ عناصرها من $\mathscr{A}$ على النحو الآتي:

$$B_i = \begin{cases} A_i & for\ \ i \leq n \\ \varnothing & for\ \ i > n \end{cases}$$

عندئذ يكون لدينا ما يلي:

$$\bigcup_{i=1}^{\infty} B_i = A_1 \cup A_2 \cup \ldots \cup A_n \cup \varnothing \cup \ldots\ldots \cup \varnothing \cup \ldots = \bigcup_{i=1}^{n} A_i \in \mathscr{A}$$

وهذا يعني أنَّ $\mathscr{A}$ مغلق بالنسبة إلى عملية الاتحاد غير المنتهي العدود أيضاً، وبهذا يتمّ الإثبات.

### أمثلة (١٠ , ٢ , ١ , ٤)

لتكن $\Omega$ فضاء الحوادث الابتدائية لتجربة عشوائية ما، فعندئذ:

١- لنأخذ صف الحوادث $\mathscr{A} = \{\varnothing , \Omega\}$، فنجد أنَّ $\mathscr{A}$ هو جبر فوق $\Omega$. إنَّ هذا الجبر يُعرَف باسم **الجبر المبتذل** Trivial Algebra فوق $\Omega$، وهو أضعف جبر فوق $\Omega$، بمعنى أنَّه إذا وجد جبر آخر $\mathscr{B}$ فوق $\Omega$ فإنَّه سيكون لدينا $\mathscr{A} \subseteq \mathscr{B}$. كذلك نجد أنَّ صف الحوادث $\mathscr{A}$ هو $\sigma$- **جبر** فوق $\Omega$ أيضاً، ويحمل الاسم السّابق نفسه، وهو أضعف $\sigma$- **جبر** فوق $\Omega$ أيضاً.

٢- من أجل حادث $A \subset \Omega$ لنأخذ صف الحوادث $\mathscr{A} = \{\varnothing , A , \bar{A} , \Omega\}$، فنجد أنَّ هذا الصف يُشكِّل جبراً فوق $\Omega$ (وكذلك يُشكِّل $\sigma$- **جبر** فوق $\Omega$ أيضاً). إنَّ هذا النوع من الجبور يُدعى **الجبر المولَّد بالحادث** $A$.

٣- إذا كان $\Omega$ منتهياً أو غير منته ولكن قابل للعدّ على الأكثر، وأخذنا $\mathscr{A}=2^{\Omega}$، فعندئذ سيكون هذا الصف $\sigma$- **جبر** فوق $\Omega$، وهو أكبر $\sigma$- **جبر** فوق $\Omega$ أيضاً، بمعنىً أنَّه إذا كان $\mathscr{B}$ هو $\sigma$- **جبر** فوق $\Omega$، فإنَّه سيكون $\mathscr{B} \subseteq \mathscr{A}$، ولذلك يُقال عن $\mathscr{A}$ إنَّه **أقوى جبر** Strongest Algebra أو **أغنى جبر** Richest Algebra فوق $\Omega$ (علماً أنَّ هذه التسمية الأخيرة هي اجتهاد وليست من التسميات المتعارف عليها كما للجبر المبتذل).

٤- إذا كانت $\Omega$ مجموعة نتائج تجربة إلقاء حجر نرد لمرة واحدة، فعندئذ سيكون لدينا $\Omega = \{1,2,3,4,5,6\}$، ولنأخذ صفوف المجموعات الآتية:

$$\mathscr{A}_1 = \{\varnothing , \Omega , \{1,3,5\}, \{2,6\}, \{2,5\}, \{1,3,4,6\}, \{1,2,3,5,6\}\}$$

$$\mathscr{A}_2 = \begin{Bmatrix} \varnothing, \Omega, \{1,3,5\}, \{2,4,6\}, \{2,5\}, \{1,3,4,6\}, \{1,2,3,5\}, \\ \{1,3,4,5,6\}, \{2,4,5,6\}, \{1,2,4,5,6\}, \{3\}, \{1,3\}, \{2\} \end{Bmatrix}$$

$$\mathscr{A}_3 = \{ \varnothing, \Omega, \{1,5\}, \{2,3,4,6\}, \{2,5\}, \{1,3,4,6\} \}$$

$$\mathscr{A}_4 = \{ \varnothing, \Omega, \{1\}, \{2,3,4,5,6\} \}$$

فنلاحظ أنَّ الصف $\mathscr{A}_4$ هو الصف الوحيد الذي يحقِّق شروط الجبر فوق $\Omega$، ويُمثِّل الجبر المولَّد بالحادث $A = \{1\}$.

سنعرض فيما يلي إحدى الطرائق المستخدمة في توليد الجبور فوق $\Omega$ وذلك بشكل مُبسَّط، ومن أجل ذلك لنعرِّف أولاً التجزئة لمجموعة ما $E$ غير خالية.

### (١١ , ٢ , ١ , ٤) تعريف (التجزئة لمجموعة Partition of a Set)

لتكن $E$ مجموعة ما غير خالية و $I \subseteq \mathbb{N}$ (**لاحظ أنَّ $I$ مجموعة قابلة للعدّ على الأكثر**)، ولنأخذ أسرة المجموعات:

$$\mathscr{Z} = \{ Z_i \subseteq E \ ; i \in I \}$$

فعندئذ يُقال عن $\mathscr{Z}$ إنَّها تجزئة لـ $E$ **إذا وفقط إذا** تحقَّقت البنود الآتية:

**a)** $\quad Z_i \neq \varnothing \quad ; \forall\, i \in I$

**b)** $\quad Z_i \bigcap_{i \neq j} Z_j = \varnothing \quad ; \forall\, i, j \in I$

**c)** $\quad E = \bigcup_{i \in I} Z_i$

إنَّ المجموعات $Z_i$ المكوِّنة للتجزئة تُسمَّى **ذرات التجزئة** $\mathscr{Z}$ Atoms of the Partition.

### (١٢ , ٢ , ١ , ٤) ملاحظات

١- في الحالة الخاصة عندما تكون $\mathscr{Z} = \{E\}$، فعندئذ يُقال عن $\mathscr{Z}$ إنَّها **التجزئة المبتذلة** Trivial Partition للمجموعة $E$.

٢- إذا كانت $E = \Omega$ مجموعة نتائج تجربة عشوائية، وكانت:

$$\mathscr{Z} = \{ Z_i \ ; Z_i \subseteq \Omega \ , i \in I \}$$

تجزئة لـ $\Omega$ مع $Z_i$ حوادث من أجل كل $i \in I$، فإنَّ $\mathscr{Z}$ تُدعى **صفاً تاماً** من الحوادث.

٣- إذا كانت $\Omega$ مجموعة نتائج تجربة عشوائية، وكانت $\Omega$ منتهية أو غير منتهية ولكنَّها قابلة للعدّ على الأكثر، فعندئذ:

أ- سنأخذ $\mathscr{A} = 2^{\Omega}$ على أنَّه الـ $\sigma$- **جبر** فوق $\Omega$ ما لم نشير إلى خلاف ذلك أو يُفرض علينا نموذجاً معيَّناً من الجبور نتيجة للمسألة التي قيد الدراسة.

ب- يمكننا وبشكل وحيد أن نشكِّل جبراً $\mathscr{A}$ فوق $\Omega$ على النحو الآتي:

$$\mathscr{A} = \left\{ \varnothing \ , \ \sum_{i=1}^{n} Z_i \quad ; n \in \mathbb{N}_{|I|} \right\}$$

وهذا الجبر $\mathscr{A}$ يُدعى **الجبر المولَّد** من التجزئة $\mathscr{Z}$، ويُكتب حينئذ $\mathscr{A} = \alpha(\mathscr{Z})$.

في الواقع، إنَّ الطرح السّابق قابل للعكس بمعنى أنَّه إذا كان $\mathscr{A}$ جبراً فوق $\Omega$ فإنَّه توجد تجزئة وحيدة $\mathscr{Z}$ للمجموعة $\Omega$ بحيث تكون ذراتها من $\mathscr{A}$، ويكون لدينا في هذه الحالة $\mathscr{A} = \alpha(\mathscr{Z})$.

**خلاصة القول:**

تحت الفرضيات السابقة يوجد تقابل واحد لواحد بين أسرة كل الجبور فوق $\Omega$ وأسرة كل التجزئات للمجموعة $\Omega$.

**(٤، ١، ٢، ١٣) أمثلة**

١- الجبر المبتذل $\mathscr{A} = \{\varnothing, \Omega\}$ يولّد من التجزئة المبتذلة $\mathscr{Z} = \{\Omega\}$.

٢- الجبر المولّد من حادث $A \subset \Omega$ يولّد من التجزئة $\mathscr{Z} = \{A, \bar{A}\}$.

٣- إذا كانت $\Omega = \{\omega_1, \omega_2, ..., \omega_n\}$ منتهية، فإنَّ الجبر القوي $2^{\Omega}$ يولّد من التجزئة $\mathscr{Z} = \{\{\omega_1\}, \{\omega_2\}, ..., \{\omega_n\}\}$.

٤- إذا كانت $\Omega = \{\omega_1, \omega_2, ..., \omega_n, ...\}$ غير منتهية ولكنّها قابلة للعدّ، فإنَّ الجبر القوي $2^{\Omega}$ يولّد من التجزئة الآتية:

$$\mathscr{Z} = \{\{\omega_1\}, \{\omega_2\}, ..., \{\omega_n\}, ...\}$$

يُقدّم الجدول الآتي قراءةً توضّح أوجه الاختلاف بين نظرية المجموعات ونظرية الاحتمالات في تفسير بعض الكائنات الرياضياتية التي تعاملنا معها سابقاً.

الجدول (٤، ٣)

| الرمز | التفسير وفقاً لنظرية المجموعات | التفسير وفقاً لنظرية الاحتمالات |
|---|---|---|
| $\omega$ | عنصر أو نقطة | حادث ابتدائي أو نتيجة تجربة عشوائية |
| $\Omega$ | مجموعة عناصر | مجموعة نتائج تجربة عشوائية أو فضاء الحوادث الابتدائية أو الحادث الأكيد |
| $\varnothing$ | المجموعة الخالية | الحادث المستحيل |
| $\mathscr{A}$ | صف مجموعات، وعندما يكون مُحقّقاً لشروط الجبر هو جبر المجموعات | صف مجموعات، وعندما يكون مُحقّقاً لشروط الجبر هو جبر لحوادث |
| $A \in \mathscr{A}$ | عنصر من صف مجموعات | حادث |
| $\bar{A}$ | المجموعة المتمّمة لمجموعة $A$ بالنسبة إلى مجموعة شاملة $\Omega$ | الحادث الذي سيتحقّق في حال عدم تحقّق الحادث $A$ من فضاء الحوادث الابتدائية $\Omega$ |
| $A \cup B$ | مجموعة العناصر التي تنتمي إلى المجموعة $A$ أو إلى $B$ أو إلى كليهما | الحادث الذي يتحقّق بتحقّق الحادث $A$ أو الحادث $B$ أو تحقّق كليهما معاً |
| $A \cap B$ | مجموعة العناصر التي تنتمي إلى كلٍّ من المجموعتين $A$ و $B$ معاً | الحادث الذي يتحقّق بتحقّق الحادثين $A$ و $B$ في آنٍ واحد |
| $A \setminus B$ | مجموعة العناصر التي تنتمي إلى المجموعة $A$ ولا تنتمي إلى المجموعة $B$ | الحادث الذي يتحقّق بتحقُّق الحادث $A$ ولكن دون تحقُّق الحادث $B$ |
| $A \Delta B$ | مجموعة العناصر التي تنتمي إلى مجموعة $A$ أو إلى $B$ ولكن دون أن تنتمي إلى كليهما | الحادث الذي يتحقّق بتحقّق الحادث $A$ أو تحقّق الحادث $B$ ولكن دون تحقُّقهما معاً |

**(٤، ١، ٣) تعيين الدالّة الاحتمالية:**

يُعدُّ تعيين الدالّة الاحتمالية من أصعب المراحل الثلاث الخاصّة بتعيين الفضاء الاحتمالي لتجربة عشوائية، وذلك لأنَّ المرحلة الأولى المتمثّلة بتعيين فضاء الحوادث الابتدائية $\Omega$، والثانية المتمثّلة بتعيين جبر الحوادث فوق $\Omega$، ليستا بالأمر الصعب. أمّا الصعوبة في تعيين الدالّة الاحتمالية فتكمن في أكثر من جانب منها:

١- كيف يُمكننا تعيين احتمال الحادث الابتدائي في تجربة عشوائية ما؟

٢- ما هي الخصائص التي يجب على الدالّة تحقيقها حتى تصبح عديمة التناقض لدى تعميمها؟

في الحقيقة يمكننا الجزم بأنَّ الإجابة على السؤال الأول ليس من صلب نظرية الاحتمالات، وذلك لأنَّ احتمال حادث ابتدائي متعلّق بطبيعة المادة التي تجرى عليها التجربة، فعلى سبيل المثال:

أ) لو أخذنا تجربة قذف قطعة نقود متوازنة لمرّة واحدة فقط، فعندئذ سيكون نسبة ظهور (هذه النسبة سيصبح لها بعد قليل اسم الاحتمال)

كل حادث ابتدائي متعلّق بهذه التجربة يساوي $\frac{1}{2} = \frac{1}{|\Omega|}$، ولكن في حال ليس لدينا معلومات حول تجانس المادة التي صنِّعت منها قطعة النقود، فعندئذ لا يمكننا الادعاء أنَّ نسبة ظهور كل حادث ابتدائي متعلّق بهذه التجربة يساوي $\frac{1}{2}$، فقد يكون مركز ثقل القطعة منحازاً إلى أحد وجهي القطعة بسبب عدم تجانس المادة المكوِّنة للقطعة، ومن ثَّم يكون للوجه الآخر نسبة أكبر في الظهور للأعلى.

ب) لو أخذنا تجربة قذف قطعة نقود متوازنة لمرّتين متتاليتين، فعندئذ سيكون نسبة ظهور كل حادث ابتدائي متعلّق بهذه التجربة يساوي $\frac{1}{4} = \frac{1}{|\Omega|}$.

ج) لو أخذنا تجربة قذف قطعة نقود متوازنة لـ $n$ مرّة متتالية، فعندئذ سيكون نسبة ظهور كل حادث ابتدائي متعلّق بهذه التجربة يساوي $\frac{1}{2^n} = \frac{1}{|\Omega|}$.

أمّا للإجابة على السؤال الثاني فقد بُذلَت من أجله محاولات جادة من قبل بعض علماء الرياضيات (وعلماء الاحتمالات على وجه الخصوص)، وقد تراوحت نتائجهم ما بين عدم الدقّة حيناً والتخصيص حيناً آخر، إلى أن جاء التعميم في النصف الأول من القرن العشرين (في عام **1933**) على يد الرياضياتي الروسي كلموغوراف.

فيما يلي سنقوم بعرض مُبسَّط وموجز لبعض الطرائق التي قُدِّمت لتعيين الدالّة الاحتمالية مع التنويه بعيوب كل من هذه الطرائق.

### (١ ,٣ ,١ ,٤) تعيين الدالّة الاحتمالية اعتماداً على التكرار النسبي لحادث

لنفترض أنّنا نقوم بتنفيذ تجربة عشوائية مجموعة نتائجها $\Omega$، وسنرصد (أو نراقب) من خلال هذه التجربة تحقُّق حادث ما $A$ متعلّق بهذه التجربة العشوائية، ثمّ نكرِّر هذه التجربة وتحت الشروط نفسها عدداً من المرات قدره $n$، ولنفترض أنَّ الحادث $A$ قد تحقّق $m$ مرّة خلال الـ $n$ تجربة، فعندئذ يسمَّى العدد $m$ **بـ التكرار المطلق** Absolut Frequency **للحادث $A$** ويُرمز له بـ $H_n(A)$، في حين أنَّ النسبة $\frac{m}{n}$ تُدعى **التكرار النِّسبي** Relative Frequency **للحادث $A$** ويُرمز له بـ $h_n(A)$، أي إنَّ:

$$h_n(A) = \frac{m}{n} \quad [4,2]$$

وهكذا نلاحظ أنَّ $H_n(A)$ هو عدد من $\mathbb{N}_n$، في حين أنَّ $h_n(A)$ هو عدد من المجموعة $\left\{0, \frac{1}{n}, \frac{2}{n}, ..., 1\right\}$، وعلاوة على ذلك يمكن التَحقُّق وبسهولة من صحة البنود الآتية:

١- لدينا $0 \le h_n(A) \le 1$.

٢- لدينا $h_n(\Omega) = 1$، وكذلك $h_n(\varnothing) = 0$ دوماً.

٣- من أجل أي حادثين $A$ و $B$ متعلّقين بهذه التجربة لدينا:

$$h_n(A \cup B) = h_n(A) + h_n(B) - h_n(A \cap B)$$

٤- من أجل أي حادثين $A$ و $B$ متعلّقين بهذه التجربة مع $A \supseteq B$ يكون لدينا $h_n(B) \ge h_n(A)$.

الآن بفرض أنَّ $n$ قيمة مثبّة، فإنّه من أجل أي حادث $A$ متعلّق بهذه التجربة يمكننا أن نلحقه (نقابله) بعدد غير سالب $h_n(A)$، ومن ثمَّ يصبح لدينا دالّة $P$ معرَّفة على أسرة حوادث من $\Omega$ (تحقّق شروط الجبر على $\Omega$) وقيمها في الفترة $[0, 1]$ كما يلي:

$$P(A) := h_n(A) \quad [4,3]$$

فعندئذ نجد أنَّ الدالّة $P$ تُحقِّق الخصائص الآتية:

١- من أجل أي حادث $A$ متعلّق بالتجربة العشوائية لدينا $0 \le P(A) \le 1$، وعلى وجه الخصوص $P(\varnothing) = 0$.

٢- من أجل أي حادثين $A$ و$B$ متنافيين متعلّقين بالتجربة العشوائية يكون:

$$P(A+B) = P(A) + P(B)$$

٣- لدينا $P(\Omega) = 1$، وذلك لأنّه من أجل أي حادث $A$ متعلّق بالتجربة العشوائية لدينا:

$$P(\bar{A}) = 1 - P(A) \quad \Rightarrow P(\Omega) = P(\bar{\varnothing}) = 1 - P(\varnothing) = 1$$

إنّ الدالّة $P$ المعرّفة آنفاً سُمّيت **دالّة احتمالية**، ولكن نلاحظ أنّ هذا التعريف للدالّة الاحتمالية ليس دقيقاً، وذلك لأنّها عُرِّفت من أجل قيمة مُثبّتة لـ $n$، إذ إنّه ما يدرينا أنّنا سنحصل على القيم لـ $h_n(A)$ نفسها لو أنّنا أعدنا التجربة عدداً أكبر من المرات؟

عند هذا الاعتراض على تعريف الدالّة $P$ قُدّم طرح بديل، وهو أن تُأخذ نهاية التكّرار النسبي لحادث $A$ عندما $n$ تسعى إلى اللانهاية كاحتمال لهذا الحادث، أي أن نضع:

$$P(A) = \lim_{n \to \infty} h_n(A) \qquad [4,4]$$

وهنا جاء الرد بالرفض أيضاً بناءً على نتائج مثال بيفون- بيرسون الآتي:

**(١, ١, ٣, ١, ٤) مثال (تجربة بيفون- بيرسون Buffon-Pearson Experement):**

لقد قام الرياضياتي والمؤرخ الطبيعي الفرنسي بوفون Georges-Louis de Buffon (1707-1788)، وكذلك الرياضياتي بيرسون بدراسة تأثير استقرارية التكّرار النسبي في مثال قذف قطعة نقود معدنية متوازنة (أي لكل من الوجهين النصيب نفسه في الظهور)، ورصدا حادث ظهور الصورة (وليكن $A$)، فكانت نتائجهما كما في الجدول الآتي:

الجدول (٤ , ٤)

| نتائج | عدد الرميات التي نفّذت | التكرار المطلق للحادث $A$ | التكرار النسبي للحادث $A$ |
|---|---|---|---|
| بيفون | 4040 | 2048 | 0.5080 |
| بيرسون | 12000 | 6019 | 0.5016 |
| بيرسون | 24000 | 12012 | 0.5005 |

فلو قمنا بدراسة تقارب متتالية التكرارات النسبيّة $(h_n(A))_{n\in\mathbb{N}}$ فإنّنا سنجدها متباعدة نحو الـ $\infty$ (ليست متقاربة)، وذلك لأنّه لو افترضنا جدلاً أنّ المتتالية $(h_n(A))_{n\in\mathbb{N}}$ متقاربة من العدد $h(A) = 0.5$ (التي تعني أنّ قطعة النقود متوازنة) لوجب أن يكون بحسب تعريف التقارب للمتتاليات العددية الاتي مُحقّقاً:

من أجل أي $\varepsilon > 0$ يوجد عدد صحيح غير سالب $N(\varepsilon)$ بحيث إنّه من أجل كل القيم $n > N(\varepsilon)$ يكون $|h_n(A) - h(A)| < \varepsilon$، وهذا غير متوافق مع المثال المُعطى، فلو أخذنا على سبيل المثال $\varepsilon = 0.00051$ لوجدنا من التجربة الأخيرة لـ بيرسون ما يلي:

$$|h_n(A) - h(A)| = |0.5005 - 0.5| = 0.0005 < \varepsilon$$

وسنجد منها أنّ $\frac{1}{\varepsilon} > \frac{10000}{5}$، ومن ثمّ يكون $N(\varepsilon) > 2000$، وهذا يعني أنّه من أجل $n = 2000 > N(\varepsilon)$ يجب أن يكون لدينا $|h_n(A) - h(A)| < \varepsilon$، ولكنّ هذه المتباينة غير مُحقّقة من أجل تجربة بيفون، وذلك لأنّه من أجل $n = 4040 > N(\varepsilon)$ لدينا:

$$|h_n(A) - h(A)| = |0.5080 - 0.5| = 0.0080 > \varepsilon = 0.00051$$

وهذا يعني تباعد المتتالية $(h_n(A))_{n\in\mathbb{N}}$، ومن ثمّ لا يجوز لنا أن نأخذ نهاية التكّرار النسبي لحادث $A$ كاحتمال لهذا الحادث $A$.

**خلاصة القول:**

إنَّ التكرّار النسبي لحادث لا يصلح لتعريف دالّة احتمالية وذلك لأنَّه تُؤخذ عليه العلتين الآتيتين:

١- إنَّ التكرّار النسبي لحادث لا يمكن أن يُولّد منها سوى متتالية منتهية بحيث إنَّه لا يمكننا ولا بشكل من الأشكال أن ننظر لتقارب متتالية التكرارات النسبية المستنتجة بمنظور تقارب المتتاليات العددية.

٢- حتى في حال تجاوزنا للتحذير السابق وقبولنا بتعريف الاحتمال وفقاً للعلاقة (٤ , ٤)، فإنَّه ليس هناك ما يضمن لنا حتمية تقارب متتالية التكرارات النسبية كما لاحظنا ذلك سابقاً.

على الرغم من ذلك نلاحظ أنَّ التكرّار النسبي لحادث $A$ يُظهر لنا استقرار قيمة احتمال الحادث، وذلك لأنَّه يعطينا قيمة تقريبية لاحتمال للحادث (وليست دقيقة)، وهذا يوصلنا إلى قناعة مفادها أنَّ احتمال وقوع حادث ما يمكن أن يُميَّز بعددوفق منهج ما.

العرض التالي للدالّة الاحتمالية قد يكون أكثر فائدة في إعطاء منهج مُحدَّد لتعيين الدالّة الاحتمالية ولكنَّه على كل الأحوال ليس بالكافي أيضاً حيث سنلاحظ ذلك جلياً من خلال عرض هذه الطريقة.

### (٢ , ٣ , ١ , ٤) تعيين الدالّة الاحتمالية اعتماداً على مبدأ لابلاس للحالات متساوية الإمكانية

لقد قام الرياضياتي الفرنسي لابلاس Pierre-Simon, marquis de Laplace (1749-1827) بمحاولة جادة لتعريف الاحتمال لحادث من خلال حالة خاصة عندما تكون مجموعة نتائج التجربة العشوائية $\Omega$ منتهية، وكذلك لجميع نتائج التجربة العشوائية النصيب نفسه في الظهور. حيث لاحظ من مواقف عمليّة أنّه يمكن الاستفادة من استخدام خاصية التجانس للمادة التي تُجرى عليها التجربة (مثل قطعة نقود مصنّعة من مادة متجانسة)، وهي تقتضي أن يصبح لجميع النتائج النصيب نفسه في الظهور.

فلو افتراضنا أنَّ $\Omega=\{\omega_1,\omega_2,\ldots,\omega_n\}$ مع $n\in\mathbb{N}$ مثبّت، فإنَّه إذا كان احتمال الحادث الابتدائي $\omega_i$ هو $p_i$ من أجل كل $i\in\mathbb{N}_n$، فعندئذ سيكون لدينا $p_1=p_2=\ldots=p_n=1/n$، ومن ثمَّ يكون احتمال أي حادث $A$ متعلّق بالتجربة العشوائية يساوي النسبة $\frac{|A|}{n}=\frac{|A|}{|\Omega|}$، وهذا يعني أنّه يمكن تعريف دالّة $P$ على أسرة حوادث متعلّقة بهذه التجربة العشوائية (تحقّق شروط الجبر على $\Omega$) من خلال العلاقة الآتية:

$$P(A)=\frac{|A|}{|\Omega|} \qquad [4,5\text{-a}]$$

وقد صاغها لابلاس على النحو الآتي أيضاً:

$$\text{احتمال الحادث الذي قيد الدراسة} = \frac{\text{عدد الحالات الملائمة لهذا الحادث}}{\text{عدد الحالات الممكنة للتجربة}}$$

ومن ثمَّ يكون:

$$P(A)=\frac{|A|\ \text{عدد الحالات الملائمة للحادث}\ A}{|\Omega|\ \text{عدد الحالات الممكنة للتجربة}} \qquad [4.5\text{-b}]$$

وهنا نلاحظ أنَّ الدالّة $P$ المُعطاة من خلال العلاقة [4.5-a] (أو [4.5-b]) تُحقّق البنود الآتية:

١- من أجل أي حادث $A$ متعلّق بالتجربة العشوائية يكون $0\le P(A)\le 1$، وعلى وجه الخصوص لدينا $P(\varnothing)=0$ دوماً.

٢- من أجل أي حادثين $A$ و $B$ متنافيين متعلّقين بالتجربة العشوائية يكون لدينا $P(A+B)=P(A)+P(B)$.

٣- لدينا $P(\Omega)=1$.

وقد أُطلق على $P$ اسم **الدالّة الاحتمالية**، والعلاقة الأخيرة [4.5-b] التي وضعها لابلاس أُطلق عليها اسم **مبدأ لابلاس للحالات المتساوية الإمكانية**، وعُرفت فيما بعد باسم **التعريف التقليدي للاحتمال** Classical Definition of Probability.

إنَّ التعامل مع هذا التعريف للاحتمال سيقود المرء في كثير من الحالات إلى استخدام التحليل التوافقي (**استخدام التباديل، التراتيب والتوافيق**) لحل المسائل، ومن ثمَّ سنلاحظ أنَّ التحليل التوافقي سيقوم بدور مهمّ في الحساب الاحتمالي لمسائل تتوافق وحالتنا هذه. هذا من جانب، ومن جانب آخر يجب على المرء ألّا يتجاوز فرضيات هذا التعريف إذ إنَّه يصبح عديم الفائدة عندما يكون فضاء الحوادث الابتدائية غير منته (**وذلك لإنَّ احتمال أي حادث سوف يسعى إلى الصفر**) أو عندما يكون للحوادث الابتدائية احتمالات مختلفة.

### (١, ٢, ٣, ١, ٤) أمثلة

١- لنفرض أنَّه لدينا صندوقان **I** و**II** يحويان كرات متماثلة بألوان مختلفة، وبحيث يحوي الصندوق الأول **I** على كرتين بيضاء اللون وكرة سوداء بينما يحوي الصندوق الثاني **II** كرة بيضاء وكرة سوداء وأخرى زرقاء. الآن نقوم بخلط الكرات في الصندوق الأول جيداً ومن ثمَّ نسحب (سحب عشوائي) كرة من هذا الصندوق ونضعها في الصندوق الثاني دون رؤيتها، وبعد ذلك نقوم بخلط الكرات في الصندوق الثاني جيداً ومن ثمَّ نسحب كرة من هذا الصندوق. عندئذ لنحسب احتمال أن تكون الكرة المسحوبة أخيراً ذات لون أسود.

الحل: نلاحظ أولاً أنَّ لجميع الحوادث الابتدائية الناتجة عن هذه التجربة النصيب نفسه في الظهور، وكذلك فضاء الحوادث الابتدائية لهذه التجربة منته، فلو رمزنا للكرتين البيضاويين في الصندوق الأول بـ $W_{11}$ و $W_{12}$ وللكرة السوداء في هذا الصندوق بـ $B_{11}$ بينما سنرمز للكرة البيضاء والسوداء والزرقاء في الصندوق الثاني على الترتيب بـ $W_{21}$ , $B_{21}$ , $BL$ ، فعندئذ يكون لدينا:

$$\Omega=\begin{Bmatrix}(W_{11},W_{11}),(W_{11},W_{21}),(W_{11},B_{21}),(W_{11},BL),(W_{12},W_{12}),(W_{12},W_{21}),\\(W_{12},B_{21}),(W_{12},BL),(B_{11},B_{11}),(B_{11},B_{21}),(B_{11},W_{21}),(B_{11},BL)\end{Bmatrix}$$

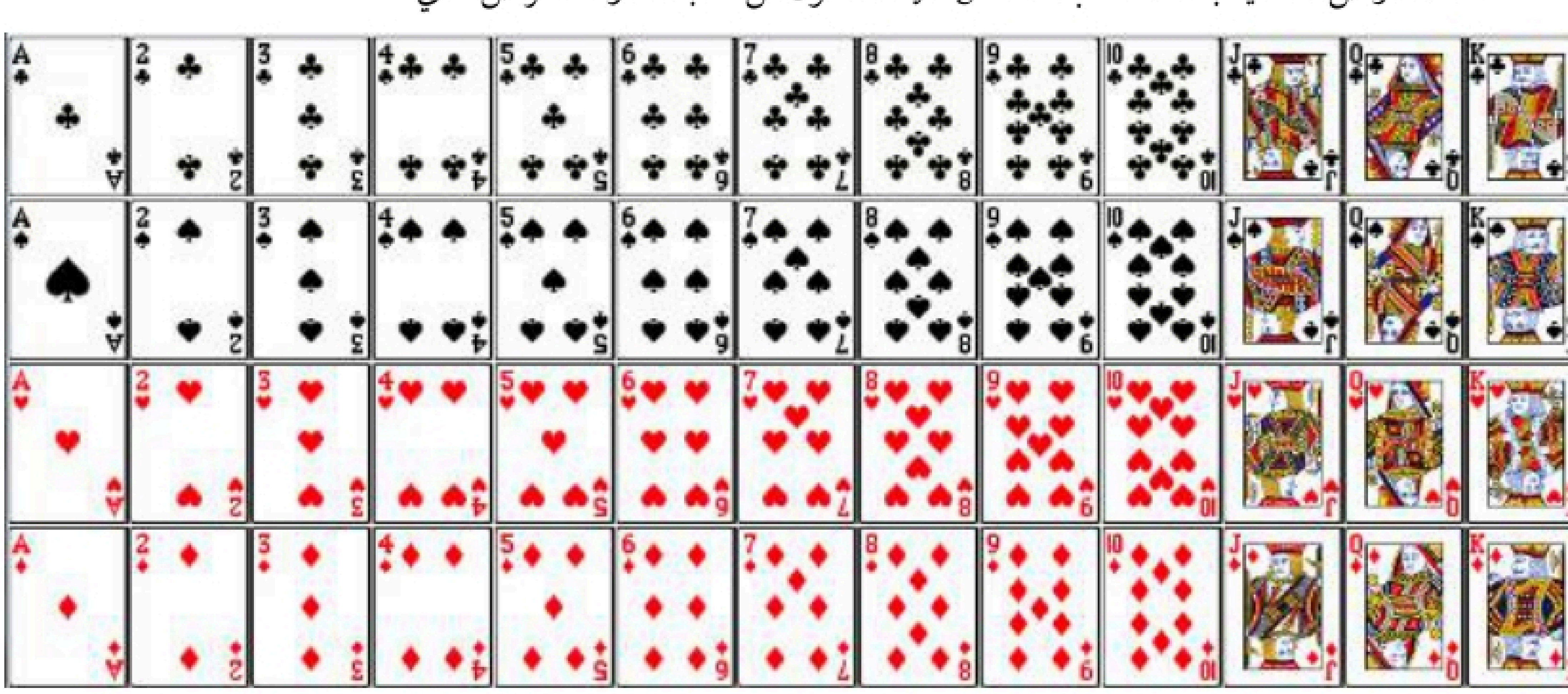
حيث ترمز المركبة الأولى من الثنائية إلى نتيجة السحب الأول بينما ترمز المركبة الثانية من الثنائية إلى نتيجة السحب الثاني، وبفرض أنَّ $A$ هو حادث سحب كرة سوداء من الصندوق الثاني بَعد إضافة كرة إليه من الصندوق الأول فإنَّه يمكننا أن نكتب:

$$A=\{(W_{11},B_{21}),(B_{11},B_{11}),(W_{12},B_{21}),(B_{11},B_{21})\}$$

ومن ثمَّ بحسب مبدأ لابلاس في الاحتمالات يكون لدينا:

$$P(A)=|A|/|\Omega|=4/12=1/3$$

٢- لنفترض أنه لدينا بطاقات لعب Playing Carts مكوّن من 52 بطاقة ولها العرض الآتي:



**لدينا أربعة أنواع من البطاقات هي:** ♦ ،♥، ♠ و ♣ **وتُدعى دينار** Diamond، **قلب** Heart، **بستوني** Spade و **زهر** Club **على الترتيب**.


قمنا بخلط البطاقات جيداً ومن ثمَّ سحبنا ست بطاقات دفعة واحدة، فما هو احتمال أن نحصل على العرض الجانبي لهذه البطاقات؟

**الحل**: نعلم أنَّه يوجد أربع بطاقات من كل نموذج ولونان وأربعة أصناف في هذا النوع من البطاقات. عندئذ يمكننا اختيار بطاقة من هذه البطاقات المماثلة لها بعدد من الطرائق يساوي:

$$\binom{4}{1} = \frac{4!}{1!\cdot 3!} = 4$$

ومن جهة أخرى بحسب المبدأ الأساسي في العدّ (قاعدة الضرب- انظر الملحق A-) يمكننا اختيار هذه البطاقات الست من بين مثيلاتها بعدد من الطرائق يساوي:

$$\binom{4}{1}\cdot\binom{4}{1}\cdot\binom{4}{1}\cdot\binom{4}{1}\cdot\binom{4}{1}\cdot\binom{4}{1} = \left(\binom{4}{1}\right)^6 = 4^6 = 4096$$

ومنه بفرض أنَّ $A$ هو حادث الحصول على البطاقات الست المعروضة سابقاً، فإنَّ عدد الحالات الملائمة لهذا الحادث يساوي 4096، وأمّا عدد الحالات الممكنة للتجربة، وهو عدد الطرائق لسحب 6 بطاقات دفعة واحدة من 52 بطاقة، فإنَّه يساوي:

$$\binom{52}{6} = \frac{52!}{6!\cdot 46!} = 20358520$$

ومنه بحسب مبدأ لابلاس في الاحتمالات يكون للاحتمال المطلوب القيمة الآتية:

$$P(A) = \frac{4096}{20358520} = 0.000201$$

٣- لنفترض أنه لدينا بطاقات لعب مكوَّنة من 52 بطاقة، قمنا بخلط البطاقات جيداً ومن ثمَّ سحبنا ثلاث بطاقات في آنٍ واحد، فما هو احتمال أن تكون البطاقات الثلاث:

١- جميعها واحدات ♠ Aces؟

٢- جميعها ألماسات ♦ diamonds؟

٣- من أصناف مختلفة؟

٤- من ألوان مختلفة؟

**الحـــل**: من أجل الإجابة على هذه الأسئلة لنرمز بـ :

$A$ لحادث الحصول على ثلاث واحدات لدى عملية السحب.

$B$ لحادث الحصول على بطاقات ألماس لدى عملية السحب.

$C$ لحادث الحصول على بطاقات من أنواع مختلفة لدى عملية السحب.

$D$ لحادث الحصول على بطاقات من ألوان مختلفة لدى عملية السحب.

عندئذ يكون لدينا:

$$P(A) = \frac{|A|}{|\Omega|} = \binom{4}{3}\Big/\binom{52}{3} = 0.00018 \qquad \& \qquad P(B) = \frac{|B|}{|\Omega|} = \binom{13}{3}\Big/\binom{52}{3} = 0.0129$$

$$P(C) = \frac{|C|}{|\Omega|} = 4\binom{13}{1}\cdot\binom{13}{1}\cdot\binom{13}{1}\Big/\binom{52}{3} = 0.3976 \qquad \& \qquad P(D) = P(\varnothing) = 0$$

وذلك لأنَّه لدينا لونان فقط فإنَّ $D$ لا تحوي أية نتيجة، أي إنَّ $D = \varnothing$ .

### (٣,٣,١,٤) تعيين الدالَّة الاحتمالية هندسياً

لقد لاحظنا أنَّ التعريف التقليدي للاحتمال قابل للتطبيق في حالات خاصة فقط، ولذلك حاول بعض الرياضياتين تقديم تعريف الاحتمال وفق عرض هندسي من خلال تمثيل الحوادث بأشكال هندسية مناسبة، اعتقاداً منهم أنَّهم قد تجاوزوا السلبيات التي اعترضت طريقة التكرّار النسبي والتعريف التقليدي للاحتمال فكانت الانطلاقة على النحو الآتي:

إنَّ أية نتيجة للتجربة العشوائية يمكن أن تُفَسَّر (ومن أسس هندسية) على أنَّها قذف عشوائي لنقطة على مجموعة أساس $\Omega$ من الفضاء الإقليدي $\mathbb{R}^n$، وهنا يجب أن يُفْهَم من كلمة عشوائي ما يلي:

١- إنَّ النقطة المقذوفة يمكن لها أن تقع (تنطبق) على أي نقطة من $\Omega$.

٢- الحوادث التي لها القياس نفسه (الطول نفسه - المساحة نفسها - الحجم نفسه ...) سيكون لها النصيب نفسه في الظهور.

عندئذٍ يُحسب احتمال حادث متعلّق بالتجربة العشوائية من خلال العلاقة الآتية:

$$\text{احتمال الحادث الذي قيد الدراسة} = \frac{\text{قياس الشكل الممثّل لهذا الحادث}}{\text{قياس } \Omega}$$

ومن ثَمَّ يمكن حساب احتمال حادث $A$ متعلّق بالتجربة من خلال العلاقة الآتية:

$$P(\Omega) = \frac{\text{قياس الشكل الممثّل لهذا الحادث } A}{\text{قياس } \Omega} = \frac{m(A)}{m(\Omega)} \qquad [4,6]$$

وهذه العلاقة اشتهرت باسم التعريف الهندسي للاحتمال Geometric Definition of Probability.

### (١,٣,٣,١,٤) ملاحظات

١- إنَّ الحوادث التي لها حيّز ذو بُعد أصغر من بُعد الحيّز الأساس $\Omega$ يكون لها قياس يساوي الصفر (قياسها معدوم)، ومن أمثال هذه الحوادث: نقطة في مستقيم - مستقيم في مستوٍ - مستوٍ في الفراغ الثلاثي وما فوق أو...

٢- إنَّ احتمال حادث $A$ يكون مستقلاً عن الطول والشكل الخاص بهذا الحادث في $\Omega$، وذلك لأنَّ احتمال هذا الحادث هو نسبة قياس هذا الحيّز إلى قياس $\Omega$. بمعنى آخر، إنَّ احتمال حادث $A$ من $\Omega$ يكون مساوياً لسلوك النسبة بين الحيِّز الملائم للحادث $A$ وبين الحيِّز الذي يشغله $\Omega$.

٣- نلاحظ أنَّ الدالَّة $P$ المُعطاة من بالعلاقة [4.6] تُحقّق البنود الآتية:

أ- من أجل أي حادث $A$ متعلّق بالتجربة العشوائية لدينا $0 \le P(A) \le 1$، وعلى وجه الخصوص $P(\varnothing) = 0$ دوماً، علماً أنَّ الحادث المستحيل $\varnothing$ هنا هو ذلك الحادث الذي لا يحوي أية نتيجة من نتائج التجربة العشوائية وبُعد الحيّز الذي يشغله يساوي بُعد الحيّز الأساس $\Omega$ أيضاً.

ب من أجل أي حادثين $A$ و$B$ متنافيين متعلّقين بهذه التجربة يكون لدينا $P(A+B) = P(A) + P(B)$.

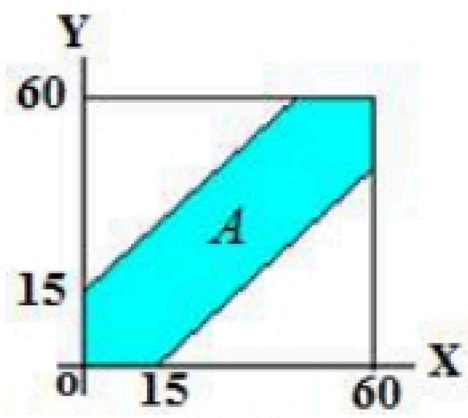
٤- إنَّ الحوادث التي لها حيّز ذو بُعد أصغر من بُعد الحيّز الأساس $\Omega$ يكون لها قياس يساوي الصفر كما ذكرنا ذلك سابقاً، ولكنَّها ليست حوادث مستحيلة، وتسمَّى حوادث شبه مستحيلة Almost Impossible Events.

بناءً على ما سبق تمَّ التعامل مع الدالَّة $P$ كدالَّة احتمالية، وهنا نلاحظ أنَّ هذا العرض للحساب الاحتمالي يبدي انسجاماً وتطابقاً مع التعريف التقليدي للاحتمال إذا ما أخذنا بالحسبان أنَّ الحوادث التي لها حيّز مشغول بالقياس نفسه يكون لها الاحتمال نفسه في الظهور (يقابل في التعريف التقليدي للاحتمال أن لجميع الحوادث الابتدائية النصيب نفسه في الظهور)، إلاّ أنَّ التعميم في هذه الطريقة يكمن في أنَّ

عدد الحوادث الابتدائية يمكن أن يكون غير منته.

المثال الآتي يُعدُّ من الأمثلة النموذجية على الحساب الاحتمالي باستخدام المفهوم الهندسي للاحتمال.

### (٢, ٣, ٣, ١, ٤) مثال التقاء شخصين



الشكل (١ , ٤)

اتفق شخصان على الالتقاء في مكان مُحدَّد وخلال مُدَّة ساعة معيَّنة بحيث إنَّه إذا قدم أحدهما إلى المكان المحدَّد خلال هذا الموعد ولم يجد زميله فإنَّه ينتظر لمدة 15 دقيقة فقط، ثمَّ يغادر المَكان مباشرة، والسؤال المطروح هو: **ما هو احتمال التقاء الشخصين؟**

**الحل**: من أجل ذلك لنأخذ الجملة المتعامدة **XoY** حيث المحور **oX** يرمز للزمن الخاص بالشخص الأول، وأمَّا **oY** فإنَّه يرمز للزمن الخاص بالشخص الآخر، فيكون ل $\Omega$ الحيِّز المعيَّن بالمربع $[0\,,60]\times[0\,,60]$ والموضَّح في الشكل الجانبي آنفاً، أي إنَّ $\Omega=[0\,,60]\times[0\,,60]$، ومن ثمَّ يكون لدينا:

$$m(\Omega)=60\times 60=3600$$

وأمَّا الحادث $A$ الذي يعبِّر عن التقاء الشخصين فهو الحيِّز:

$$A=\left\{(x\,,y)\in\mathbb{R}^2 \;\mid\; 0<x<60 \;\;\&\;\; 0<y<60 \;\;with\; |x-y|\le 15\right\}$$

وقياس الحيِّز الذي يشغله يساوي:

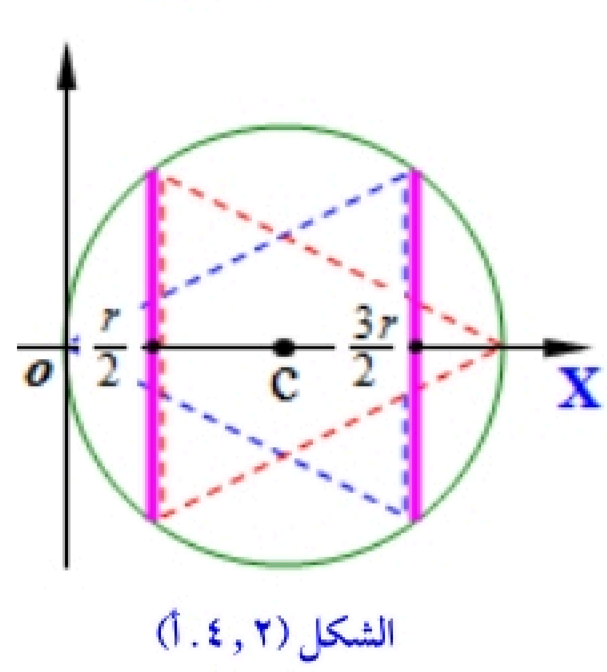
$$m(A)=m(\Omega)-2\frac{(45)(45)}{2}=1575$$

ومن ثمَّ ينتج لدينا أنَّ الاحتمال المطلوب يساوي:

$$P(A)=\frac{m(A)}{m(\Omega)}=0.4375$$

في هذا الجانب نشير إلى أنَّ التعريف الهندسي للاحتمال قد لقي في بدايات تقديمه أنواعاً من سوء الفهم والخطأ في الاستخدام، وإلى النقد اللاذع الذي دفع بعضهم إلى رفض الحساب الاحتمالي كفرع من الفروع العلمية، وذلك عندما قام البعض بعرض بعض المسائل التي نتج عنها حلول مختلفة وفقاً لطرائق مختلفة للحل على حدّ زعمهم، حيث يجب أن تكون نتيجة الحل ثابتة بغض النظر عن الطريقة المتبعة في الحل. في الحقيقة إنَّ هذه النتائج المختلفة لحلول بعض المسائل لا يعود في حقيقة الأمر إلى تناقض في المفهوم الهندسي للاحتمال، وإنَّما يعود إلى عدم وجود ضوابط دقيقة للمسائل المطروحة آنذاك، فعلى سبيل المثال قام الرياضياتي والفيلسوف البريطاني **برتراند** Arthur William Russell Bertrand (1872-1970) بتقديم مثال في عام 1889، و يُعْرَف هذا المثال باسم **مُحيِّرة برتراند**.

### (٣, ٣, ٣, ١, ٤) مسألة (مُحيِّرة برتراند Bertrand Paradox)



الشكل (٢ , ٤ . أ)

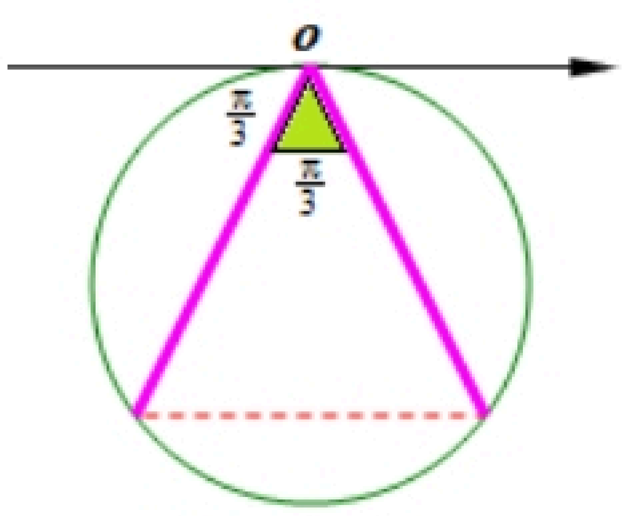
نقوم بسحب عشوائي لوتر في دائرة نصف قطرها $r$، فما هو احتمال أن يكون طول هذا الوتر أكبر أو يساوي طول ضلع مثلث متساوي الأضلاع مرسوم داخل هذه الدائرة؟

لقد قدَّم **برتراند** ثلاث طرائق لحل هذه المسألة هي كالآتي:

**الطريقة الأولى**: نعطي توجيهاً للوتر، وسنفترض أنَّ الوتر سيتحرك بشكل عمودي على قطر هذه الدائرة (انظر الشكل الجانبي)، فعندئذبفرض $A$ هو الحادث الذي يعبِّر عن الحالة التي يكون فيها طول الوتر أكبر أو يساوي طولَ ضلع مثلث متساوي الأضلاع مرسوم داخل هذه الدائرة، فإنَّ هذا الحادث يوافق الحيِّز المحصور بالمستقيمين $x=\frac{3\,r}{2}$ و $x=\frac{r}{2}$، ومن ثمَّ ينتج لدينا:

$$P(A)=\frac{m(A)}{m(\Omega)}=\frac{r}{2r}=\frac{1}{2}$$

وهكذا نلاحظ أنه فُسِّر مفهوم السحب العشوائي على أنَّه قذف نقطة على الفترة $[0,2r]$، وهذا يعني أنَّه لدينا $\Omega=[0,2r]$ .

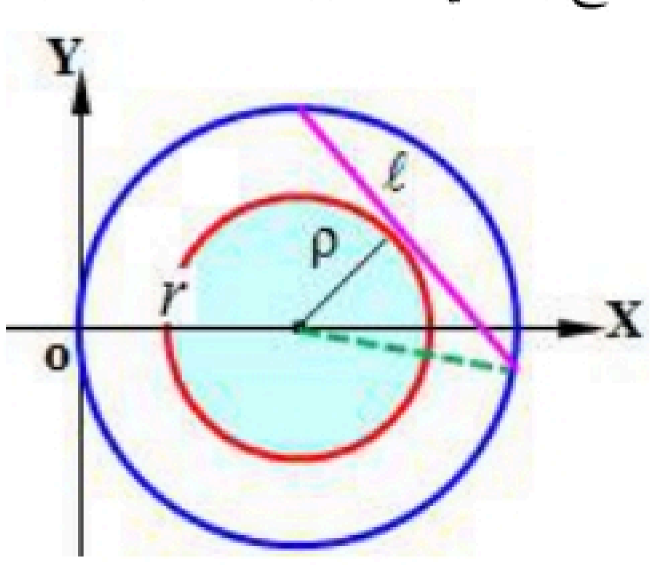


شكل (٢، ٤. ب)

**الطريقة الثانية**: في هذه الطريقة ثبِّت إحدى نهايتي الوتر على طرف الدائرة وترك الوتر يتأرجح بشكل عشوائي داخل هذه الدائرة (انظر الشكل الجانبي)، فعندئذ يكون الحّيز الموافق للحادث $A$ هو الحيِّز الموافق لزاوية محيطية قدرها $\frac{\pi}{3}$ ومجاورة لزاوية قدرها $\frac{\pi}{3}$ مع المماس للدائرة عند نقطة تثبيت الوتر مع الدائر، وهكذا ينتج لدينا:

$$P(A)=\frac{m(A)}{m(\Omega)}=\frac{\frac{\pi}{3}}{\pi}=\frac{1}{3}$$

فلاحظ هنا أنَّه فُسر مفهوم السحب العشوائي على أنَّه قذف نقطة على الفترة $[0,\pi]$، ومن ثَّم أصبح لدينا في هذه الحالة $\Omega=[0,\pi]$ .



شكل (٢، ٤. ج)

**الطريقة الثالثة**: من المعلوم أنَّ الوتر يمكن أن يعيّن وبشكل وحيد من خلال نقطة منتصف الوتر، ومن ثَّم بفرض أنَّ $\rho$ هو بُعد الوتر عن مركز الدائرة و $\ell$ هو طول الوتر (انظر الشكل الجانبي)، فإنَّه سيكون لدينا $\ell=2\sqrt{r^2-\rho^2}$ ، ومن ثَّم الحادث $A$ يَتحقَّق إذا كان $\sqrt{3}\cdot r\leq\ell$ ، وهذا يعني أنَّ الحادث $A$ سيتحقَّق إذا كان $\rho\leq 0.5\,r$ ، ومن ثَّم يكون لدينا:

$$P(A)=\frac{m(A)}{m(\Omega)}=\frac{\pi\rho^2}{\pi r^2}=\frac{\pi\left(\frac{r}{2}\right)^2}{\pi r^2}=\frac{1}{4}$$

وفي هذه الطريقة نلاحظ أنَّه فُسِّر السحب العشوائي للوتر على أنَّه قذف نقطة على سطح دائرة نصف قطرها $r$، وهذا يعني أنَّ مجموعة نتائج التجربة هي:

$$\Omega=\left\{(x\ ,\ y)\in\mathbb{R}^2\mid x^2+y^2\leq r\right\}$$

هكذا نلاحظ أنَّه حصل على ثلاث قيم لاحتمال الحادث المفروض $A$، وأنَّ كل طريقة أعطت قيمة مختلفة عن الأخرى، إنَّ السبب في ذلك يعود لتفسيرات مختلفة لمفهوم السحب العشوائي للوتر وليس بسبب خلل في التعريف الهندسي للاحتمال. إذ إنَّه فُسِّر السحب العشوائي في الطريقة الأولى للحل على أنَّه قذف نقطة على فترة طولها $2r$ وكانت مجموعة نتائج التجربة العشوائية وفقاً لهذا الطريقة هي $\Omega=[0,2r]$ ، وأمَّا في الطريقة الثانية للحل فقد فُسِّر السحب العشوائي للوتر على أنَّه قذف نقطة على فترة طولها $\pi$ وكانت مجموعة نتائج التجربة العشوائية وفقاً لهذا الطريقة هي $\Omega=[0,\pi]$ ، وأخيراً فُسِّر السحب العشوائي للوتر في الطريقة الثالثة للحل على أنَّه قذف نقطة على سطح دائرة نصف قطرها $r$ وكانت مجموعة نتائج التجربة العشوائية وفقاً لهذه الطريقة هي:

$$\Omega=\left\{(x\ ,\ y)\in\mathbb{R}^2\mid x^2+y^2\leq r\right\}$$

وبذلك نجد أنَّه لم يُحدَّد مفهوم السحب العشوائي للوتر بشكل تام وتُرك ليُستخدم دون ضوابط، ومن ثَّم فإنَّ الحلول الثلاثة السابقة ليست حلولاً للمسألة المطروحة، وإنَّما هي حلول لثلاث مسائل مختلف بعضها عن بعض الآخر، وأنَّ المسألة المطروحة نفسها هي مسألة غير قابلة للحل ما لم يُحدَّد وبدقّة ما يعنيه بالسحب العشوائي للوتر (كيف سيتم سحب الوتر؟).

### (٤ ,٣ ,٣ ,١ ,٤) عيوب التعريف الهندسي للاحتمال

إذن، وكما لاحظنا فإنَّ استخدام المفهوم الهندسي للاحتمال ليس بالأمر السهل ويخفي في جنباته العديد من العيوب، ولكن وفقاً لمعايير شديدة تعتمد في جانب كبير منها على نظرية القياس أمكن توليد فرع من نظرية الاحتمالات يُعرف باسم **الهندسة العشوائية** Stochastic Geometry، ولهذا العلم تطبيقات عمليّة واسعة في مجال تقنية الاتصالات والتصوير الفضائي (الاستطلاع). أمّا أبرز عيوب المفهوم الهندسي للاحتمال فهي:

١- صعوبة إيجاد ضوابط دقيقة لمفهوم العشوائية في مجال الهندسة.

٢- صعوبة تعيين الشكل الهندسي الموافق للحادث.

٣- صعوبة العمليات الحسابية باستخدام الهندسة في الحالة العامّة.

**خلاصة القول:**

وجدنا ممّا سبق أنَّ كل طريقة من الطرائق السّابقة المستخدمة في تعيين الدالَّة الاحتمالية تعاني من عيوب ومساوئ لا يمكن تجنبها وفقاً للأسس المتَّبعة في هذا التعيين، فعلى سبيل المثال يصادفنا في كثير من الأحيان تجارب عشوائية ذات عدد غير منته من النتائج، لا بل قد تكون مجموعة نتائج التجربة غير قابلة للعدّ أيضاً، كأن تكون $\Omega$ مساوية لفترة من $\mathbb{R}$ أو لمجال من $\mathbb{R}^n$، وفي هذه الحالة لا يمكننا النظر إلى جميع عناصر $2^{\Omega}$ على أنّها حوادث، إذ إنَّه أُثبت وجود مجموعات من $2^{\mathbb{R}}$ لا يمكن النظر إليها كحوادث، ومنها على سبيل المثال لا الحصر المجموعات غير البوريلية. كذلك من الممكن أن يكون لدينا حوادث مُركَّبة لعدد غير منته من الحوادث، ومن ثمَّ مفهوم الجبر يصبح غير ذي جدوى ونكون مضطرين للتعامل مع مفهوم الـ $\sigma$- **جبر** . لذلك كان من الضرورة البحث في منهج شمولي لتعيين الدالَّة الاحتمالية بحيث يُجنّبنا التناقضات التي يمكن أن تظهر لنا عند التعميم، وهذا يعني إعطاء الدالَّة الاحتمالية شروطاً مُحدّدة تستجيب لمتطلبات الحالات العامّة في الدراسات الاحتمالية التي يمكن أن تصادفنا، وهذا ما تقدمه لنا الفقرة التالية.

### (٤ ,٣ ,١ ,٤) تعيين الدالَّة الاحتمالية وفقاً لمنهج رياضياتي مجرد

من أجل إعطاء الدالَّة الاحتمالية مميِّزات شمولية (عامّة) وضعت أسس رياضياتية مُجّردة مبنيَّة على مجموعة من المسلَّمات لبناء الفضاء الاحتمالي من أجل أيّة تجربة عشوائية، وهـذه المسلَّمات هي:

**المسلّمة الأولى**: مجموعة نتائـج التجربة العشـوائية $\Omega$ ليست خالية، أي إنَّ $\Omega \neq \varnothing$، وهذه الخاصية مثبتة دوماً.

**المسلّمة الثانية**: صف الحوادث المعتمد $\mathscr{A} \subseteq 2^{\Omega}$ غير خالٍ وبحيث يكون من نوع الـ $\sigma$- **جبر** فوق $\Omega$.

**المسلّمة الثالثة**: الدالَّة الاحتمالية $P$ تُعرَّف على $\mathscr{A}$ كدالَّة حقيقية غير سالبة، وبحيث تُحقِّق ما يلي:

١- يجب أن يكون $P(\varnothing) = 0$ . إنَّ هذا الشرط (**والذي هو الشرط الأول من تعريف القياس Measure- في نظرية القياس-**) يعني أنَّ الدالَّة الاحتمالية $P$ ليست مبتذلة، بمعنى أنّه يوجد على الأقل حادث $A \in \mathscr{A}$ بحيث يكون $P(A) > 0$، ولكن يجب الانتباه هنا بأنَّ هذا الشرط لا يعني أنَّ الحادث المستحيل هو الحادث الوحيد في $\mathscr{A}$ الذي من أجله $P(\varnothing) = 0$، إذ إنَّه من الممكن أن يوجد حادث $A \in \mathscr{A}$ أو أكثر من أجله لدينا $P(A) = 0$ أيضاً، وأمثال هذه الحوادث تُدعى **حوادث شبه مستحيلة** (كما ذكرنا ذلك سابقاً).

٢- من أجل أي متتالية حوادث $(A_n)_{n\in\mathbb{N}}$ من $\mathscr{A}$ عناصرها متنافية مثنى مثنى (أي إنَّ $A_i \bigcap\limits_{i\neq j} A_j = \varnothing$) يكون لدينا:

$$P\left(\sum_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n) \qquad [4,7]$$

وهذا الشرط الأخير يوضِّح لنا بجلاء ضرورة أن يكون جبر الحوادث $\mathscr{A}$ من نوع $\sigma$- **جبر** ولا يكفي أن يكون جبراً فقط.

٣- يجب أن يكون $P(\Omega)=1$، وهنا يجب الانتباه إلى أنَّ هذا الشرط لا يعني أنَّ الحادث الأكيد هو الحادث الوحيد في $\mathscr{A}$ الذي من أجله $P(\Omega)=1$، إذ إنَّه من الممكن أن يوجد حادث $\Omega \neq A \in \mathscr{A}$ من أجله لدينا $P(A)=1$ أيضاً، وأمثال هذا الحادث $A$ تُدعى **حوادث شبه أكيدة** Almost Certainly Events.

**(١ ,٤ ,٣ ,١ ,٤) ملاحظات:**

١- إنَّ الثنائية $[\Omega , \mathscr{A}]$ التي عناصرها مُحقَّقة للمسلَّمتين الأولى والثانية تُدعى **فضاءً مقيساً** (**أو فضاءً قيوساً أو فضاءً قابلاً للقياس**) Measurable Space.

٢- إنَّ الدالَّة $P$ المُحقِّقة للشرطين الأول والثاني من المسلَّمة الثالثة تُدعى **قياساً** فوق (أو على) الفضاء المقيس $[\Omega , \mathscr{A}]$.

٣- إنَّ القياس $P$ والمُحقِّق للشرط الثالث من المسلَّمة الثالثة يُدعى **قياساً احتمالياً** على الفضاء المقيس $[\Omega , \mathscr{A}]$.

٤- إنَّ الثلاثية $[\Omega, \mathscr{A}, P]$ التي عناصرها مُحقِّقة للمسلَّمات الثلاث السابقة تُدعى **فضاءً احتمالياً** Probability Space.

٥- في الحالة الخاصة عندما تكون $\Omega=\mathbb{R}^n$ مع $n \in \mathbb{N}$ مثبَّت و $\mathbb{R}^1=\mathbb{R}$ فإنَّنا سنرمز للـ σ- **جبر** على $\mathbb{R}^n$ بالرمز $\mathfrak{R}^n$ مع $\mathfrak{R} := \mathfrak{R}^1$ الذي يُطلق عليه اسم **الحقل البوريلي** Borel Field، وأمَّا عناصر $\mathfrak{R}^n$ فيُطلق عليها اسم **مجموعات بوريلية** Borel Sets، ويُنظر إليها كحوادث.

٦- في الحالة الخاصة عندما تكون $\Omega=\{\omega_1,\omega_2,\dots,\omega_n\}$ مجموعة منتهية، فإنَّه يمكننا تعيين الدالَّة الاحتمالية $P$ انطلاقاً من معرفة احتمالات الحوادث الابتدائية، وذلك من خلال تعريف دالَّة حقيقية غير سالبة $Q$ على أسرة كل الحوادث البسيطة في $\mathscr{A}$ كما يلي:

$$Q \ :\{\{\omega_1\},\{\omega_2\},\dots,\{\omega_n\}\} \longrightarrow [0,\infty)$$
$$\{\omega_i\} \mapsto Q(\{\omega_i\}) := p_i \qquad ; i \in \mathbb{N}_n$$

أي إنَّ $p_i$ هو احتمال الحادث الابتدائي $\omega_i$ الذي يُفترض أن يكون معلوماً وبحيث يكون $\sum_{i=1}^{n} p_i = 1$، وبعد ذلك يمكننا أن نُعرِّف احتمال أي حادث $A$ من خلال العلاقة الآتية:

$$P(A)=\sum_{\omega \in A} Q(\{\omega\}) = \sum_{\substack{i=1 \\ \omega_i \in A}}^{n} Q(\{\omega_i\}) \qquad [4,8]$$

تطبيقات على هذه الطريقة تجدها في الأمثلة (٥ ,٤ ,١ ,٤) القادمة.

**(٤ ,١ ,٤) خصائص الدوال الاحتمالية Characteristics of Probability Functions**

إنَّ المبرهنات الآتية التي سنقدِّمها دون برهان تعرض لنا بعض الخصائص البسيطة للدالَّة الاحتمالية، ولكنَّها مهمَّة في إطار دراستنا هذه.

**(١ ,٤ ,١ ,٤) مبرهنة**

ليكن $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً، فعندئذ:

١- من أجل أي حادث $A$ من $\mathscr{A}$ يكون لدينا $P(\bar{A})=1-P(A)$.

٢- من أجل أي حادثين $A$ و $B$ من $\mathscr{A}$ مع $A \subset B$ سيكون $P(B) \geq P(A)$، وهذه الخاصية تُعرَف باسم **خاصية الاطراد** Monotone Property للقياس الاحتمالي $P$، وعلاوةً على ذلك يكون لدينا:

$$P(B \setminus A)=P(B)-P(A)$$

٣- من أجل أية متتالية منتهية من الحوادث $A_1$، $A_2$، ... و $A_n$ عناصرها من $\mathscr{A}$ يكون:

$$P\left(\bigcup_{k=1}^{n} A_k\right)=\sum_{k=1}^{n} P(A_k)-\sum_{i<j} P(A_i \cap A_j)+\sum_{i<j<k} P(A_i \cap A_j \cap A_k) \pm \ldots +(-1)^{n-1} P(A_1 \cap A_2 \cap \ldots \cap A_n) \qquad [4,9]$$

والذي يُدعى **قانون الجمع** Addition Law في الاحتمالات، حيث نجد:

أ- من أجل $n=2$ يصبح للعلاقة السابقة [4,9] العرض الآتي:

$$P(A_1 \cup A_2)=P(A_1)+P(A_2)-P(A_1 \cap A_2)$$

ب- من أجل $n=3$ يصبح للعلاقة السابقة [4,9] العرض الآتي:

$$P(A_1 \cup A_2 \cup A_3)=P(A_1)+P(A_2)+P(A_3)-P(A_1 \cap A_2)-P(A_1 \cap A_3)-P(A_2 \cap A_3)+P(A_1 \cap A_2 \cap A_3)$$

٤- لتكن $(A_n)_{n\in\mathbb{N}}$ متتالية غير منتهية من الحوادث عناصرها من $\mathscr{A}$ ، فعندئذ:

أ- إذا كانت $A_1 \supseteq A_2 \supseteq A_3 \supseteq \ldots$ (**أي إنَّ المتتالية المفترضة غير متناقصة**) فإنَّه سيكون لدينا:

$$P\left(\bigcup_{i=1}^{\infty} A_n\right)=\lim_{n\to\infty} P(A_n)$$

وتُعرَف هذه خاصية باسم **الاستمرار من الأدنى** Continuous from Below للقياس الاحتمالي $P$.

ب- إذا كانت $A_1 \subseteq A_2 \subseteq A_3 \subseteq \ldots$ (**أي إنَّ المتتالية المفترضة غير متزايدة**) فإنَّه سيكون لدينا:

$$P\left(\bigcap_{n=1}^{\infty} A_n\right)=\lim_{n\to\infty} P(A_n)$$

وتُعرَف هذه خاصية باسم **الاستمرار من الأعلى** Continuous from Above للقياس الاحتمالي $P$، أي إنَّ القياس الاحتمالي $P$ يُحقِّق خاصية الاستمرار من الأعلى أيضاً.

إذاً، فالقياس الاحتمالي $P$ يحقِّق خاصيتي الاستمرار من الأدنى والأعلى، ولذلك يُقال عنه إنَّه **مستمر** Continuous.

### (٢ ,٤ ,١ ,٤) تعريف قياس ديراك Dirac Measure

ليكن $\Omega$ فضاء الحوادث الابتدائية لتجربة عشوائية ما، و $\mathscr{A}$ هو $\sigma$- **جبر** فوق $\Omega$. الآن من أجل $\omega \in \Omega$ سنأخذ $\delta_\omega$ دالّة حقيقية مُعرَّفة على $\mathscr{A}$ كما يلي:

$$\mathscr{A} \ni A \mapsto \delta_\omega(A)=\begin{cases}0 & for \quad \omega \notin A\\ 1 & for \quad \omega \in A\end{cases} \qquad [4,10]$$

ويدعى $\delta_\omega(\bullet)$ **قياس ديراك** المُركَّز على $\omega$ (**نسبة إلى الفيزيائي النظري الإنكليزي ديراك Paul Adrien Maurice Dirac (1902 – 1984)**).

في الواقع يلعب هذا القياس دوراً مُهمًّا جداً في صياغة العلاقات للقياسات الاحتمالية في التوزيعات المتقطِّعة، وهذا ما تظهره لنا المبرهنة الآتية.

### (٣ ,٤ ,١ ,٤) مبرهنة

ليكن $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً، فإذا كان فضاء الحوادث الابتدائية $\Omega$ منتهياً أو غير منتهٍ ولكن قابل للعدّ، فعندئذٍ سيكون للقياس الاحتمالي $P$ العرض الآتي:

$$P(A)=\sum_{\omega\in\Omega} P(\{\omega\})\,\delta_\omega(A) \quad ;\forall A\in\mathscr{A} \qquad [4,11]$$

علماً أنَّ $P(\{\omega\})$ هو احتمال الحادث الابتدائي $\omega$ الذي يجب أن يكون معلوماً.

### (٤، ١، ٤، ٤) ملاحظات

١- من تعريف قياس ديراك نلاحظ أنَّه من أجل كل $\omega\in\Omega$ ما يلي محقَّقاً:

أ) لدينا $\delta_\omega(\varnothing)=0$.

ب) من أجل أي متتالية حوادث $(A_n)_{n\in\mathbb{N}}$ من $\mathscr{A}$ عناصرها متنافية مثنى مثنى يكون لدينا:

$$\delta_\omega\left(\sum_{n=1}^{\infty}A_n\right)=\sum_{n=1}^{\infty}\delta_\omega(A_n)$$

ج) لدينا $\delta_\omega(\Omega)=1$.

وبالتالي من أجل أي $\omega\in\Omega$ مثَّبت لدينا $[\Omega,\mathscr{A},\delta_\omega]$ هو فضاء احتمالي أيضاً.

٢- بفرض أنَّ فضاء الحوادث الابتدائية $\Omega$ منته أو غير منته ولكن قابل للعدّ، فعندئذ يكون $\sum_{\omega\in\Omega}P(\{\omega\})=1$، ولإثبات ذلك سنفترض أنَّ $\mathscr{Z}=\{Z_i\ ;i\in I,\ I\ finite\}$ تجزئة للحادث الأكيد $\Omega$، فعندئذ يمكننا أن نكتب الآتي:

$$\begin{aligned}P(\Omega)&=P\left(\sum_{i\in I}Z_i\right)=\sum_{i\in I}P(Z_i)=\sum_{i\in I}\ \sum_{\omega\in\Omega}P(\{\omega\})\,\delta_\omega(Z_i)=\sum_{\omega\in\Omega}P(\{\omega\})\sum_{i\in I}\delta_\omega(Z_i)\\&=\sum_{\omega\in\Omega}P(\{\omega\})\,\delta_\omega\left(\sum_{i\in I}Z_i\right)=\sum_{\omega\in\Omega}P(\{\omega\})\,\delta_\omega(\Omega)=\sum_{\omega\in\Omega}P(\{\omega\})\end{aligned}$$

وبما أنّ $P(\Omega)=1$ فإنَّه ينتج صحة دوانا.

فيما يلي سنقدِّم بعض الأمثلة البسيطة على كيفية تعيين الفضاء الاحتمالي وفقاً للمسلّمات الثلاث الخاصّة بالفضاء الاحتمالي.

### (٤، ١، ٤، ٥) أمثلة

١- لنأخذ تجربة إلقاء قطعة نقود متوازنة لثلاث مرَّات متتالية، ولنقم بتعيين الفضاء الاحتمالي $[\Omega,\mathscr{A},P]$ لهذه التجربة العشوائية، ومن ثمَّ حساب احتمال الحصول على صورتين على الأقل من هذه التجربة.

الحل: لتعيين الفضاء الاحتمالي $[\Omega,\mathscr{A},P]$ لهذه التجربة العشوائية علينا تعيين كل مركبة من مركباته، حيث لدينا:

أ- مجموعة نتائج هذه التجربة العشوائية هي:

$$\Omega=\left\{\underbrace{HHH}_{\omega_1},\overbrace{HHT}^{\omega_2},\underbrace{HTH}_{\omega_3},\overbrace{THH}^{\omega_4},\underbrace{TTH}_{\omega_5},\overbrace{THT}^{\omega_6},\underbrace{TTH}_{\omega_7},\overbrace{TTT}^{\omega_8}\right\}$$

ب- بما أنَّ $\Omega$ منتهية فإنَّه يمكننا أخذ الـ $\sigma$ - جبر $\mathscr{A}=2^\Omega$، وهو أقوى جبر على $\Omega$.

ج- بما أنَّ قطعة النقود متوازنة فإنه سيكون لجميع النتائج النصيب نفسه في الظهور، ومن ثمَّ يصبح لدينا:

$$Q(\{\omega_i\})=p_i=\frac{1}{8} \qquad ;\forall i=\mathbb{N}_8$$

ومن ثَمّ يمكننا أن نُعرِّف احتمال أي حادث $A\in\mathscr{A}$ من خلال العلاقة الآتية:

$$P(A)=\sum_{i\,:\,\omega_i\in A} Q(\{\omega_i\})$$

الآن بفرض أنَّ $A$ هو حادث الحصول على صورتين على الأقل، فإنَّه سيكون احتمال $A$ يساوي:

$$P(A)=\sum_{i=1;\,\omega_i\in A}^{8} Q(\{\omega_i\})=Q(\{\omega_1\})+Q(\{\omega_2\})+Q(\{\omega_3\})+Q(\{\omega_4\})=\frac{4}{8}=\frac{1}{2}$$

وأمّا إذا استخدمنا المبرهنة السابقة (٣ , ٤ , ١ , ٤) فإنَّه سيكون لدينا:

$$\begin{aligned}P(A)=\sum_{\omega\in\Omega} P(\{\omega\})\delta_\omega(A) &= P(\{\omega_1\})\cdot 1+P(\{\omega_2\})\cdot 1+P(\{\omega_3\})\cdot 1+P(\{\omega_4\})\cdot 1+P(\{\omega_5\})\cdot 0\\ &\quad +P(\{\omega_6\})\cdot 0+P(\{\omega_7\})\cdot 0+P(\{\omega_8\})\cdot 0\\ &=\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}=\frac{4}{8}=\frac{1}{2}\end{aligned}$$

علما أنَّه لدينا $P(\{\omega\})=\frac{1}{8}$ من أجل كل $\omega\in\Omega$ .

٢- لنأخذ تجربة إلقاء حجر نرد متوازن لمرَّة واحدة فقط، ولنقم بتعيين الفضاء الاحتمالي $[\Omega,\mathscr{A},P]$ لهذه التجربة العشوائية، ومن ثمَّ حساب احتمال الحصول على عدد فردي من هذه التجربة.

الحـل: من أجل تعيين الفضاء الاحتمالي $[\Omega,\mathscr{A},P]$ لهذه التجربة العشوائية نجد ما يلي:

أ- مجموعة نتائج تجربة إلقاء حجر نرد لمرَّة واحدة هي $\Omega=\{1,2,3,4,5,6\}$ .

ب- بما أنَّ $\Omega$ منتهية فإنَّه يمكننا أخذ الـ $\sigma$- **جبر** فوق $\Omega$ هو $\mathscr{A}=2^\Omega$ .

ج- بسبب أنَّ $\Omega$ منتهية فإنَّه يمكننا أن نُعرِّف احتمال أي حادث $A\in\mathscr{A}$ من خلال العلاقة الآتية:

$$P(A)=\sum_{i\,;\,\omega_i\in A} Q(\{\omega_i\})$$

وبما أنَّ حجر النرد متوازن فإنَّه سيكون لدينا:

$$Q(\{\omega_i\})=p_i=\frac{1}{6}\qquad ;i=1,2,3,4,5,6$$

والآن بفرض أنَّ $A$ هو حادث الحصول على عدد فردي، فإنَّه سيكون لدينا احتمال الحصول على عدد فردي من هذه التجربة يساوي:

$$P(A)=\sum_{i=1;\,\omega_i\in A}^{6} Q(\{\omega_i\})=Q(\{\omega_1\})+Q(\{\omega_3\})+Q(\{\omega_5\})=\frac{3}{6}=\frac{1}{2}$$

وأمّا باستخدام المبرهنة السابقة (٣ , ٤ , ١ , ٤) فإنَّنا نجد:

$$\begin{aligned}P(A)&=\sum_{\omega\in\Omega} P(\{\omega\})\delta_\omega(A)\\ &=P(\{\omega_1\})\cdot 1+P(\{\omega_2\})\cdot 0+P(\{\omega_3\})\cdot 1+P(\{\omega_4\})\cdot 0+P(\{\omega_5\})\cdot 1+P(\{\omega_6\})\cdot 0\\ &=\frac{1}{6}+\frac{1}{6}+\frac{1}{6}=\frac{3}{6}=\frac{1}{2}\end{aligned}$$

علماً أنَّه لدينا في هذا المثال $P(\{\omega\})=\frac{1}{8}$ من أجل كل $\omega$ من $\Omega$.

## (٤,٢) الاحتمالات الشرطية

**Conditional Probability**

لقد قمنا فيما سبق بالتعرّف على بنية الفضاء الاحتمالي لتجربة عشوائية ما، وحساب احتمالات لحوادث تتعلّق بتجربة عشوائية مفروضة، إلا أنَّه يصادفنا في كثير من الحالات حساب احتمالات لحوادث ذات طبيعة شرطية، فعلى سبيل المثال لو أمعنا النظر في الأسئلة الآتية:

أ- ما هو احتمال هطول الأمطار علماً أنَّ السماء مُلبَّدةً بالغيوم؟

ب- ما هو احتمال وقوع حادث سير على طريق معيَّن علماً أنَّه طريق زلق؟

ج- في مصنع لإنتاج سلعة معينة يوجد خطوط للإنتاج $L_1، L_2$ ، ... ، $L_k$ قمنا بسحب قطعة من الإنتاج الكلي للمصنع فوجدناها تالفة، فما هو احتمال أن تكون هذه القطعة من إنتاج الخط $L_2$ ؟

نلاحظ هنا أنَّ المطلوب حساب احتمال حادث معيَّن إذا عُلم تحقُّق وقوع حادث آخر سابق له، ولتوضيح ذلك أكثر لنأخذ على سبيل المثال تجربة إلقاء حجر نرد متوازن لمرَّة واحدة فقط، ولنطرح السؤال الآتي:

**ما هو احتمال الحصول على الرقم 3 علماً أنّنا قد حصلنا على عدد فردي؟**

للإجابة على هذا السؤال لنرمز بـ $A$ لحادث الحصول على العدد 3 وكذلك بـ $B$ لحادث الحصول على عدد فردي، فيكون لـدينا:

$$A = \{3\} \quad \& \quad B = \{1, 3, 5\}$$

ولكن بعد علمنا بالحصول على عدد فردي فإنَّ الحالات الممكنة للتجربة ستصبح 1، 3 و5، ومن ثمَّ اخْتُزِلَ فضاء نتائج التجربة $\Omega$ إلى الحادث $B$، ومنه بحسب مبدأ لابلاس في الاحتمالات سيكون الاحتمال المطلوب يساوي النسبة الآتية:

$$\frac{\text{عدد الحالات الملائمة للحادث } A}{\text{عدد الحالات الممكنة للتجربة}} = \frac{|A \cap B|}{|B|}$$

فلو قمنا الآن بتقسيم البسط والمقام على $|\Omega|$ فإنَّه يصبح لدينا:

$$\frac{\frac{1}{6}}{\frac{3}{6}} = \frac{\frac{|A \cap B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{P(A \cap B)}{P(B)}$$

إذن احتمال الحصول على العدد 3 علماً أنّنا حصلنا على عدد فردي يساوي قيمة النسبة $\frac{P(A \cap B)}{P(B)}$ ويساوي $\frac{1}{3}$، وهنا يُشترط أن يكون $0 < P(B)$ وذلك لأنَّه لا معنى للنسبة السابقة عندما يكون $P(B) = 0$ .

الآن وبعد هذا التقديم يمكننا أن نصيغ تعريف الاحتمال الشرطي على النحو الآتي.

### (٤, ٢, ١) تعريف (الاحتمال الشرطي لحادث Conditional Probability of an Event)

ليكن $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً مُعطى، وليكن $A$ و$B$ حادثين من $\mathscr{A}$ بحيث $0 < P(B)$ . عندئذ تُدعى قيمة النسبة العددية $\frac{P(A \cap B)}{P(B)}$ بـ **الاحتمال الشرطي** للحادث $A$ علماً أنَّ الحادث $B$ تُحقَّق وقوعه مسبقاً، ويرمز له اصطلاحاً بالرمز $P(A \mid B)$ ، أي إنَّ:

$$P(A \mid B) := \frac{P(A \cap B)}{P(B)} \qquad [4,12]$$

### (٤ , ٢ , ١ , ١) ملاحظات

١- فيما يلي وعلى سبيل التبسيط والاختصار سنقرأ $P\,(A\,|\,B)$ على النحو الآتي: **الاحتمال الشرطي لـ $A$ بالنسبة إلى $B$**.

٢- في الحالة الخاصة عندما تكون $\Omega$ منتهية ولجميع نتائجها النصيب نفسه في الظهور، فإنَّه يمكننا حساب الاحتمال الشرطي من خلال العلاقة الآتية:

$$P\,(A\,|\,B) = \frac{|\,A \cap B\,|}{|\,B\,|} \quad ;|\,B\,| > 0 \qquad [4,13]$$

٣- من أجل $B$ حادث مُثبَّت من $\mathscr{A}$ مع $0 < P\,(B)$ تكون خواص الاحتمال الشرطي $P\,(\bullet\,|\,B)$ على الفضاء المقيس $[\ \Omega \cap B\ ,\ \mathscr{A} \cap B\ ]$ هي خواص الدالَّة الاحتمالية $P\,(\bullet)$ نفسها على الفضاء المقيس $[\ \Omega\ ,\ \mathscr{A}\ ]$ ، علماً أنَّ:

$$\mathscr{A} \cap B = \{A \cap B \quad ;\forall A \in \mathscr{A}\ \}$$

ولذلك يُرمز (في بعض الأحيان) للاحتمال الشرطي $P\,(\bullet|\,B)$ بالرمز $P_B\,(\bullet)$ للدلالة على أنَّ $P_B$ هنا له خصائص القياس الاحتمالي.

٤- عندما يُكتب $P\,(AB)$ إنَّما يُقصد بذلك $P\,(A \cap B)$، ولذلك سوف نقوم بحذف رمز التقاطع $\cap$ بين الحوادث على سبيل التبسيط ونذكره عند الضرورة أو على سبيل التذكير به أحياناً.

٥- من مبدأ التناظر في استخدام الرموز يمكننا في حال $0 < P\,(A)$ أن نكتب من أجل $B \in \mathscr{A}$ الآتي:

$$P\ (B\,|\,A) := \frac{P\,(A \cap B)}{P\,(A)} \qquad [4,14]$$

### (٤ , ٢ , ١ , ٢) نتائج

ليكن $[\ \Omega, \mathscr{A}, P\ ]$ فضاءً احتمالياً مفروضاً، وليكن $A$ حادثاً من $\mathscr{A}$ مع $0 < P\,(A)$ ، فعندئذ يمكن للقارئ التحقُّق من صحة النتائج الآتية:

١- لدينا $P\,(\varnothing\,|\,A) = 0$ محقَّقةً دوماً.

٢- من أجل $B$ حادث من $\mathscr{A}$ مع $A \subseteq B$ فإنَّه سيكون $P\ (B\ |A\ ) = 1$ ، ومن ثَمَّ ينتج لدينا $P\ (\Omega\,|\,A\,) = 1$ .

٣- من أجل أي حادثين متنافيين $B$ و$C$ من $\mathscr{A}$ يكون لدينا:

$$P\ (B + C\ |A\,) = P\ (B\ |A\,) + P\ (C\ |A\,) \qquad [4,15]$$

٤- من أجل أي حادثين $B$ و$C$ من $\mathscr{A}$ يكون لدينا:

$$P\ (B\,C\ |A\,) = P\ (B\ |A\,C\,) \cdot P\ (C\ |A\,) \qquad [4,16]$$

٥- من النتيجتين ٢ و٣ السابقتين ينتج لدينا أنَّ:

$$P\left(\overline{B}\ |A\right) = 1 - P\ (B\ |A\,) \qquad [4,17]$$

٦- إذا كان $A$ و$B$ حادثين كيفيين من $\mathscr{A}$ مع $0 < P\,(A)$ و $0 < P\,(B)$ ، فعندئذ نجد من تعريف الاحتمال الشرطي أنَّ:

$$P\ (A\,B\,) = P\ (A\,) \cdot P\ (B\ |A\,) = P\,(B\,) \cdot P\,(A\ |B\,) \qquad [4,18]$$

وهذه العلاقة الأخيرة تُدعى **قانون الضرب في الاحتمالات** التي يمكن تعميمها كما في المبرهنة الآتية.

**(٤ , ٢ , ٢) مبرهنة (قانون الضرب في الاحتمالات Multiplication Law of Probability)**

ليكن $[\,\Omega, \mathscr{A}, P\,]$ فضاءً احتمالياً مُعطى، ولتكن $A_1$ ، $A_2$ ،... و $A_n$ حوادث من $\mathscr{A}$ ، فإذا كان:

$$P\,(A_1 A_2\,...\,A_{n-1}) > 0 \qquad (*)$$

فعندئذ تكون العلاقة الآتية مُحَقَّقةً:

$$\begin{aligned} P\,(A_1 A_2 ... A_n) &= P\,(A_1)\cdot P\,(A_2|A_1)\cdot P\,(A_3|A_1 A_2)\cdot ... \\ & ...\cdot P\,(A_{n-1}|A_1 A_2 ... A_{n-2})\cdot P\,(A_n|A_1 A_2 ... A_{n-1}) \end{aligned} \qquad [4,19]$$

**البرهان:** لإثبات صحة هذه العلاقة يمكننا بادئ ذي بدء وبسهولة إثبات أنَّه إذا كانت العلاقة (*) مُحَقَّقةً فإنَّه سيكون لدينا:

$$P\,(A_1 A_2\,...\,A_k) > 0 \quad ; \forall k \in \mathbb{N}_{n-1}$$

وسنثبت صحة العلاقة [4,19] بطريقة الاستقراء الرياضياتي، حيث نجد من تعريف الاحتمال الشرطي أنَّه من أجل $n = 2$ لدينا:

$$P\,(A_1 A_2) = P\,(A_1)\cdot P\,(A_2\,|\,A_1)$$

وهذا يعني أنَّ العلاقة [4,19] صحيحة من أجل $n = 2$ .

الآن سنفترض أنَّ العلاقة [4,19] صحيحة من أجل $n-1$ ولنثبت صحتها من أجل قيمة ما $n \in \mathbb{N}$ ، ومن أجل ذلك سنضع $B = A_1 A_2 ... A_{n-1}$ ، فعندئذ يكون لدينا الآتي:

$$\begin{aligned} P\,(A_1 A_2 ... A_n) &= P\,(B\,A_n) = P\,(B)\cdot P\,(A_n\,|\,B) \\ &= P\,(A_1 A_2 ... A_{n-1})\cdot P\,(A_n\,|\,A_1 A_2 ... A_{n-1}) \\ &= P\,(A_1)\cdot P\,(A_2\,|\,A_1)\cdot P\,(A_3\,|\,A_1 A_2)\,....\cdot P\,(A_{n-1}\,|\,A_1 A_2 ... A_{n-2})\cdot P\,(A_n\,|\,A_1 A_2 ... A_{n-1}) \end{aligned}$$

وهذا يعني أنَّ العلاقة [4,19] صحيحة من أجل أيّة قيمة $n \in \mathbb{N}$ ، وبذلك يتمّ البرهان.

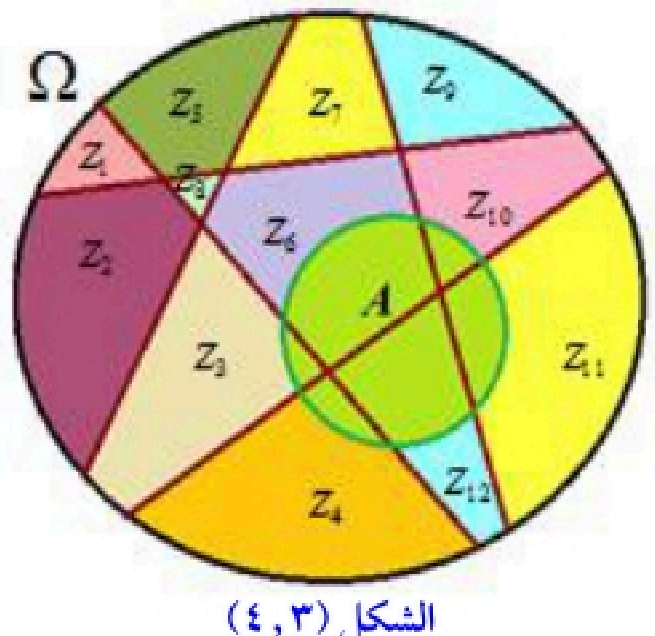
المبرهنة الآتية تمكننا من حساب احتمال حادث ما متعلِّق بالتجربة العشوائية إذا كان لدينا تجزئة للحادث الأكيد.

**(٤ , ٢ , ٣) مبرهنة (صيغة الاحتمال التام Total Probability Formula)**

ليكن $[\,\Omega, \mathscr{A}, P\,]$ فضاءً احتمالياً مُعطى، و $\mathscr{Z} = \{\,Z_i \;\; ; i \in I \subseteq \mathbb{N}\,\}$ تجزئة للحادث الأكيد $\Omega$ ذراتها من $\mathscr{A}$ وتحقِّق العلاقة $0 < P\,(Z_i)$ من أجل كل $i \in I$ ، فعندئذٍ من أجل أي حادث $A$ من $\mathscr{A}$ تكون العلاقة الآتية مُحَقَّقةً:

$$P\,(A) = \sum_{i \in I} P\,(Z_i)\cdot P\,(A\,|\,Z_i) \qquad [4,20]$$

وهذه العلاقة الأخيرة تُعرف باسم **صيغة الاحتمال التام**.




الشكل (٤ , ٣)

**البرهان:** بما أنَّ ذرات التجزئة هي حوادث متنافية مثنى مثنى فإنَّه يمكننا أن نكتب ما يلي:

$$A = A \cap \Omega = A \cap \left( \sum_{i \in I} Z_i \right) = \sum_{i \in I} A Z_i$$

وبما أنَّ $0 < P\,(Z_i)$ من أجل كل $i \in I$ ، فإنَّه باستخدام خواص القياس الاحتمالي $P$ وقانون الضرب في الاحتمالات نحصل على العلاقة الآتية التي تثبت صحة دعوانا:

$$P\,(A) = P\,(A \cap \Omega) = P\left( \sum_{i \in I} A Z_i \right) = \sum_{i \in I} P\,(A Z_i) = \sum_{i \in I} P\,(Z_i)\cdot P\,(A\,|\,Z_i)$$

الآن، بفرض أنَّ $\mathscr{Z} = \{Z_i \;\; ; i \in I \subseteq \mathbb{N}\}$ تجزئة للحادث الأكيد $\Omega$ ذراتها من $\mathscr{A}$، فإنَّه قد يصادفنا في بعض الحالات حساب الاحتمال الشرطي لذرة $Z_k$ بالنسبة إلى حادث $A \ni \mathscr{A}$ . إنَّ المبرهنة الآتية تُعطينا العلاقة التي تستخدم لحساب مثل هذا الاحتمال الشرطي، وتعرف هذه العلاقة باسم **نظرية بييز** (أو قانون بييز أو دستور بييز **Bayes' law or Bayes' rule**) في الاحتمالات، وهي تنسب إلى الإحصائي والفيلسوف الإنكليزي بييز (1701–1761) Thomas Bayes.

### (٤ , ٢ , ٤) مبرهنة (نظرية بييز Bays's theorem)

ليكن $[\, \Omega, \mathscr{A}, P \,]$ فضاءً احتمالياً مُعطى، و $\mathscr{Z} = \{Z_i \;\; ; i \in I \subseteq \mathbb{N}\}$ تجزئة للحادث الأكيد $\Omega$ ذراتها من $\mathscr{A}$ وتحقِّق العلاقة $P(Z_i) > 0$ من أجل كل $i \ni I$، فعندئذ من أجل قيمة $k \ni I$ يكون ما يلي محقَّقاً:

$$P(Z_k \mid A) = \frac{P(Z_k) \cdot P(A \mid Z_k)}{\sum_{i \in I} P(Z_i) \cdot P(A \mid Z_i)} \qquad [4,21]$$

**البرهان:** من تعريف الاحتمال الشرطي، ومن صيغة الاحتمال التام يمكننا أن نكتب الآتي من أجل كل $k \ni I$:

$$P(Z_k \mid A) = \frac{P(AZ_k)}{P(A)} = \frac{P(Z_k) \cdot P(A \mid Z_k)}{P(A)} = \frac{P(Z_k) \cdot P(A \mid Z_k)}{\sum_{i \in I} P(Z_i) \cdot P(A \mid Z_i)}$$

وبهذا يتمّ البرهان.

### (١ , ٤ , ٢ , ٤) ملاحظات

١- إنَّ العلاقة السابقة [4,21] تُعرف باسم **صيغة بييز في الاحتمالات**، وتمكننا من حساب الاحتمال الشرطي لذرة من تجزئة للحادث الأكيد إذا تحقَّق وقوع حادث ما متعلِّق بالتجربة العشوائية مسبقاً.

٢- لقد نظر بييز إلى احتمال حادث ما على أنَّه درجة اعتقاد بحدوثه أو عدم حدوثه، وهذا التعريف يسمح له أن يدرس قضايا احتمالية لا ينطبق عليها مفهوم التكّرار النسبي لأنَّه لا يُكرِّر التجربة عدداً لا منتهياً من المرّات، خاصة وأنَّه يوجد الكثير من التجارب التي لا يمكن تكّرارها، مثل احتمال الموت لكائن حي، أو الوصول إلى مجرّة ما بعيدة، أو...، ولذلك بدأ بييز باستخدام درجة الاعتقاد حول فرضية ما حيث سمّى الحوادث $Z_1$ ، $Z_2$ و... بالفرضيات المتعلّقة بالمسألة المطروحة، وأمّا درجة الاعتقاد هذه والمقابلة للاحتمال $P(Z_i)$ فأسمّاها **الاحتمال الأسبق للفرضية** $Z_i$ Prior Probability، وبعد ذلك يقوم بجمع معلومات أخرى يستخدمها ليحسن من درجة اعتقاده الذي يؤدي بدوره إلى الاحتمال $P(A \mid Z_i)$ والذي أسماه **الاحتمال اللاحق** -$a$ للفرضية $Z_i$ **a**-Superior Probability.

من الاعتراضات على هذا الطرح للاحتمال هو أنَّه وفقاً لهذا المفهوم يُعدُّ كثير من الناس ليس لهم درجة اعتقاد حول موضوعات كثيرة، ولكن أشار بييز في هذا الصدد إلى أنَّه في المواقف العمليّة وخاصّة المالية منها يكون للناس درجة اعتقاد حول بعض الفرضيات مثل الربح والخسارة والأسعار، حيث يُعدُّ ذلك كافياً كحد أدنى لاعتماد هذه الاستراتيجية في الحساب.

### (٢ , ٤ , ٢ , ٤) أمثلة

١- في مصنع لإنتاج المصابيح الكهربائية يوجد ثلاثة خطوط للإنتاج $L_1$ ، $L_2$ و $L_3$ بحيث تنتج هذه الخطوط 50%، 30% و20% على الترتيب من نسبة الإنتاج الكلي للمصنع، ونسبة المعطَّل في إنتاج هذه الخطوط على الترتيب هو 0.03، 0.02 و0.01 . الآن نقوم بسحب عشوائي لمصباح من الإنتاج الكلي للمصنع، ولنقم بحساب:

أ- احتمال أن يكون المصباح المسحوب سليماً؟

ب- احتمال أن يكون المصباح المسحوب من إنتاج الخط الأول إذا علمنا أنَّه وجد مُعطَّلاً ؟

الحـل: للإجابة على هذه الأسئلة سنفترض أنَّ $Z_1$ ، $Z_2$ و $Z_3$ هو حادث سحب مصباح من إنتاج الخط $L_1$ ، $L_2$ و $L_3$ على الترتيب، فنجد أنَّ هذه الحوادث تشكل تجزئة للحادث الأكيد (الذي يمثّل كل نتائج السحب من الإنتاج الكلي للمصنع)، ومن ثمَّ بفرض $A$ هو حادث سحب مصباح مُعطَّل من الإنتاج الكلي فعندئذ من أجل الطلب:

أ ) سنفترض أنَّ $B$ حادث سحب مصباح سليم من الإنتاج الكلي، فيكون لدينا $P(B) = 1 - P(A)$ ، ولكن من صيغة الاحتمال التام لدينا:

$$P(A) = \sum_{i=1}^{3} P(Z_i) \cdot P(A \mid Z_i) = \frac{50}{100} \cdot \frac{3}{100} + \frac{30}{100} \cdot \frac{2}{100} + \frac{20}{100} \cdot \frac{1}{100} = 0.023$$

ومن ثمَّ يكون لدينا:

$$P(B) = 1 - P(A) = 0.977$$

ب) لدينا الاحتمال المطلوب $P(Z_1 \mid A)$ ، وباستخدام صيغة بييز نجد:

$$P(Z_1 \mid A) = \frac{P(Z_1)P(A \mid Z_1)}{\sum_{i=1}^{3} P(Z_i).P(A \mid Z_i)} = \frac{(0.50)(0.03)}{0.023} = 0.6522$$

وبهذا ينتهي حل المثال.

٢- لدينا ثلاثة صناديق بحيث يحوي الصندوق الأول أربع كرات حمراء وثلاث كرات سوداء بينما يحوي الثاني ثلاث كرات حمراء وخمس كرات بيضاء، وأخيراً يحوي الصندوق الثالث ثلاث كرات سوداء وثلاث كرات بيضاء. الآن بفرض أنَّ لكلّ صندوق النصيب نفسه في السحب وأنَّ جميع الكرات متماثلة تماماً (ومن ثمَّ لها النصيب نفسه في الاختيار)، وأنَّنا قمنا بسحب عشوائي لصندوق من هذه الصناديق، ومن ثمَّ سحب كرة من هذا الصندوق، فعندئذ:

أ - ما هو احتمال أن تكون الكرة المسحوبة بيضاء ؟

ب- إذا علمنا أنَّ الكرة المسحوبة كانت بيضاء، فما هو احتمال أن تكون هذه الكرة قد سُحبت من الصندوق الثاني؟

الحـل: للإجابة على هذه الأسئلة سنفترض أنَّ $Z_1$ ، $Z_2$ و $Z_3$ هو حادث سحب كرة من الصندوق الأول والثاني والثالث على الترتيب، فنجد أنَّ $Z_1$ ، $Z_2$ و $Z_3$ تُشكِّل تجزئة للحادث الأكيد $\Omega$ (الذي يمثّل مجموعة كل السحوبات الممكنة للكرات)، وكذلك سنفترض أنَّ $W$ هو حادث سحب كرة بيضاء من الصندوق، فعندئذ من أجل الطلب:

أ- يكون الاحتمال المطلوب هو $P(W)$ ، وبما أنَّه لدينا تجزئة للحادث الأكيد، فإنَّه باستخدام صيغة الاحتمال التام نجد ما يلي:

$$P(W) = \sum_{i=1}^{3} P(Z_i) \cdot P(W \mid Z_i) = \frac{1}{3} \cdot \frac{0}{7} + \frac{1}{3} \cdot \frac{5}{8} + \frac{1}{3} \cdot \frac{3}{6} = 0.375$$

ب- يكون الاحتمال المطلوب هو $P(Z_2 \mid W)$ ، وبما أنَّ لدينا تجزئة للحادث الأكيد، فإنَّه باستخدام صيغة بييز يكون لدينا:

$$P(Z_2 \mid W) = \frac{P(Z_2) \cdot P(W \mid Z_2)}{\sum_{k=1}^{3} P(Z_k) \cdot P(W \mid Z_k)} = \frac{\frac{1}{3} \cdot \frac{5}{8}}{0.375} = 0.555$$

وبهذا ينتهي حل المثال.

٣- لنأخذ تجربة إلقاء قطعة نقود متوازنة لثلاث مرَّات متتالية، ولنقم بحساب احتمال الحصول على صورة في الرمية الأخيرة علماً أنَّنا حصلنا على صورة واحدة على الأقل خلال هذه الرميات الثلاث.

الحل: من أجل ذلك لنفترض أنَّ $A$ هو الحادث الذي يعبِّر عن حصولنا على صورة في الرمية الأخيرة و$B$ هو الحادث الذي يعبِّر عن حصولنا على صورة واحدة على الأقل خلال هذه الرميات الثلاث، فعندئذ يكون لدينا:

$$A = \{HHH, HTH, THH, TTH\}$$
$$B = \{HHH, HHT, HTH, THH, TTH, THT, HTT\}$$

ومن مُعطيات المسألة لدينا قطعة النقود متوازنة، ومنه يكون الاحتمال الشرطي لـ $A$ بالنسبة إلى $B$ يساوي:

$$P(A \mid B) = \frac{P(AB)}{P(B)} = \frac{4/8}{7/8} = \frac{4}{7}$$

وبهذا ينتهي حل المثال.

٤- لدينا وعاء يحوي $m$ كرة متماثلة منها $n$ كرة حمراء والباقي $m-n$ كرة لونها أسود، علماً أنَّ $m > n$. قمنا بسحب كرتين على التوالي دون إعادتها إلى الصندوق، فإذا علمت أنَّ الكرة المسحوبة في المرَّة الأولى كانت حمراء، فما هو احتمال أن تكون الكرة المسحوبة أخيراً سوداء؟

الحل: للإجابة على هذا السؤال سنرمز بـ $A$ لحادث سحب كرة حمراء في المرة الأولى وبـ $B$ لحادث سحب كرة سوداء في المرَّة الثانية، فعندئذ يكون $P(A) = \frac{n}{m} > 0$، ومن ثمَّ الاحتمال المطلوب يساوي:

$$P(B \mid A) = \frac{P(AB)}{P(A)} = \frac{(m-n)/(m-1)}{n/m} = \frac{m(m-n)}{n(m-1)}$$

وبهذا ينتهي حل المثال.

## (٤،٣) استقلال الحوادث

**Independence of Events**

إنَّ مفهوم الاستقلال بين الحوادث يُنظر إليه كأحد المفاهيم المهمَّة جداً في الاحتمالات، ولذلك سنقدِّم فيما يلي دراسة لهذا الموضوع، وفي فصل قادم سنبحث في استقلال المتغيِّرات العشوائية، علماً أنَّ استقلال المتغيِّرات العشوائية يردّ إلى مفهوم الاستقلال بين الحوادث أيضاً.

ليكن $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً، و$A$ و$B$ حادثين من $\mathscr{A}$، فعندئذ يُقال إنَّ الحادث $A$ مستقل Independent عن الحادث $B$ إذا كان تحقُّق وقوع الحادث $B$ لا يؤثِّر في تحقُّق وقوع الحادث $A$ ولا بأي شكل من الأشكال، وهذا يعني أنَّه بفرض $P(B) > 0$ فإنَّ $P(A \mid B) = P(A)$، ومن ثمَّ باستخدام تعريف الاحتمال الشرطي ينتج ما يلي:

$$P(AB) = P(B) \cdot P(A \mid B) = P(A) \cdot P(B)$$

ومنه ينتج لدينا أنَّه إذا كان الحادث $A$ مستقل عن الحادث $B$ فإنَّ العلاقة الآتية ستكون محقَّقةً:

$$P(AB) = P(A) \cdot P(B)$$

وبشكل مماثل يكون $B$ مستقل عن $A$ بحسب المفهوم السابق إذا تحقَّقت العلاقة السابقة أيضاً، ونقول حينئذ إنَّ الحادثين $A$ و$B$ مستقلان بعضهما عن بعض، وهكذا يمكننا أن نصيغ تعريف استقلال حادثين على النحو الآتي.

**(٤،٣،١) تعريف (استقلال حادثين)**

ليكن $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً، و $A$ و $B$ حادثين من $\mathscr{A}$ ، فعندئذ يُقال إنَّ **الحادث $A$ مستقل عن الحادث $B$** إذا تحقَّقت من أجلهما العلاقة الآتية:

$$P(A\,B) = P(A)\cdot P(B) \qquad [4,22]$$

ونلاحظ هنا أنَّ العلاقة الأخيرة [4,22] تبقى سارية المفعول حتى لو كان $P(A)=0$ أو $P(B)=0$.

الآن، وبشكل مماثل لما سبق، يمكننا صياغة الاستقلال بين ثلاثة حوادث $A_1$، $A_2$ و $A_3$ من $\mathscr{A}$ حيث يُقال عن هذه الحوادث إنَّها مستقلّة إذا من أجلها ما يلي مُحقَّقاً:

١- من أجل كل $i, j \ni \mathbb{N}_3$ مع $i \neq j$ لدينا:

$$P(A_i\,A_j) = P(A_i)\cdot P(A_j)$$

٢- لدينا:

$$P(A_1\,A_2\,A_3) = P(A_1)\cdot P(A_2)\cdot P(A_3)$$

على العموم يوجد تعريف أكثر عموميةً لاستقلال الحوادث حيث يُقدِّم صياغةً لاستقلال عددمنته من الحوادث، وهذا هو محتوى التعريف الآتي.

**(٢ , ٣ , ٤) تعريف (الاستقلال لعدد منته من الحوادث Independence of Finite Number of Events)**

ليكن $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً مُعطى، ولنأخذ $A_1$، $A_2$ و... و $A_n$ من $\mathscr{A}$، فعندئذ يُقال عن هذه الحوادث إنَّها **مستقلّة عشوائياً** Stochastic Independent (أو مستقلّة إحصائياً Statistical Independent) إذا كان من أجل أي $k \ni \mathbb{N}_n$ مع $1<k$، وكل $i_j \ni \mathbb{N}$ مع $j \ni \mathbb{N}_n$ وبحيث يكون $1 \le i_1 < i_2 < \dots < i_k \le n$، فإنَّ العلاقة الآتية مُحقَّقةً:

$$P(A_{i_1}A_{i_2}\dots A_{i_k}) = P(A_{i_1})\cdot P(A_{i_2})\cdot\dots\cdot P(A_{i_k}) \qquad [4,23]$$

**(١ , ٢ , ٣ , ٤) ملاحظات**

١- من التعريف السابق للاستقلال العشوائي يتضح لنا أنَّه إذا كانت $A_1$، $A_2$ و... و $A_n$ من $\mathscr{A}$ حوادث مستقلّة عشوائياً فإن أية أسرة جزئيّة منها ستكون مستقلّة عشوائياً أيضاً.

٢- في الحالة الخاصة عندما يكون من أجل كل $i, j \ni \mathbb{N}_n$ مع $i \neq j$ لدينا:

$$P(A_i\,A_j) = P(A_i)\cdot P(A_j) \qquad [4,24]$$

فعندئذ يُقال عن الحوادث $A_1$، $A_2$ و... و $A_n$ إنَّها **مستقلّة مثنى مثنى** Pairwise Independent، ويُلاحظ هنا أنَّه إذا كانت هذه الحوادث مستقلّة مثنى مثنى فليس بالضرورة أنَّ تكون هذه الحوادث مستقلّة عشوائياً.

٣- إذا كان من أجل حادث $A$ من $\mathscr{A}$ لدينا $P^2(A) = P(A)$، فعندئذ يُقال عن الحادث $A$ إنَّه **مستقل عن نفسه**.

**(٢ , ٢ , ٣ , ٤) نتائج**

يمكن للقارئ التَحقُّق من صحة النتيجتين الآتيتين:

١- إنَّ كلاً من الحادثين $\varnothing$ و $\Omega$ مستقل عن نفسه.

٢- إنَّ الحادث $\Omega$ مستقل عن أي حادث آخر $A$ من $\mathscr{A}$.

**(٣ , ٢ , ٣ , ٤) أمثلة**

١- لنأخذ تجربة إلقاء قطعة نقود متوازنة لثلاث مرات متتالية، وليكن $A$، $B$ و $C$ هي حوادث الحصول على صورة في الرمية الأولى، والثانية، والثالثة على الترتيب، فعندئذ نجد أنَّ هذه الحوادث مستقلّة عشوائياً، وذلك لأنَّه لدينا:

$$P(A)=P(B)=P(C)=\frac{1}{2}$$

$$P(AB)=\frac{1}{4}=P(A)\cdot P(B) \qquad \& \qquad P(AC)=\frac{1}{4}=P(A)\cdot P(C)$$

$$P(BC)=\frac{1}{4}=P(B)\cdot P(C) \qquad \& \qquad P(ABC)=\frac{1}{8}=P(A)\cdot P(B)\cdot P(C)$$

**٢**– ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً مُعطى، وبحيث:

$$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\} \qquad \& \qquad \mathscr{A}=2^{\Omega} \qquad \& \quad P(\{\omega_i\})=\frac{1}{4} \qquad ; \ \forall\, i\in\mathbb{N}_4$$

ولنأخذ الحوادث $A_1=\{\omega_1,\omega_2\}$ ، $A_2=\{\omega_1,\omega_3\}$ و $A_3=\{\omega_1,\omega_4\}$ ، فنجد أنَّ:

$$P(A_1)=P(A_2)=P(A_3)=\frac{1}{2}$$

وكذلك من أجل كل $i, j \ni \mathbb{N}_3$ مع $i \neq j$ لدينا:

$$P(A_iA_j)=P(A_i)\cdot P(A_j)=\frac{1}{4}$$

ومن ثمَّ الحوادث $A_1$ و $A_2$ و $A_3$ مستقلّة مثنى مثنى، ولكنَّنا نلاحظ أنَّها ليست مستقلّة عشوائياً وذلك لأنَّ:

$$\frac{1}{4}=P(A_1A_2A_3)\neq P(A_1)\cdot P(A_2)\cdot P(A_3)=\frac{1}{8}$$

**٣**– لنأخذ تجربة إلقاء حجر نرد متوازن لمرتين متتاليتين فقط، ولنأخذ الحوادث الآتية:

$$A=\{(i\,,\,j)\ ;i\in\mathbb{N}_6\ \&\ j=1\,,2\,,5\}$$

$$B=\{(i\,,\,j)\ ;i\in\mathbb{N}_6\ \&\ j=4\,,5\,,6\}$$

$$C=\{(i\,,\,j)\ \ ;i\,,j\in\mathbb{N}_6\ with\ i+j=9\}$$

فنجد أنَّ هذه الحوادث ليست مستقلّة مثنى مثنى، ومن ثمَّ فليست مستقلّة عشوائياً وذلك لأنَّ:

$$\frac{1}{6}=P(AB)\neq P(A)\cdot P(B)=\frac{1}{2}\cdot\frac{1}{2}=\frac{1}{4}$$

علماً أنَّه لدينا:

$$P(ABC)=\frac{1}{36}=P(A)\cdot P(B)\cdot P(C)$$

وهكذا نلاحظ من هذا المثال أنَّه إذا كانت العلاقة الآتية:

$$P(A_1A_2\ldots A_n)=P(A_1)\cdot P(A_2)\cdot\ldots\cdot P(A_n)$$

مُحقَّقةً من أجل حوادث مُعطاة $A_1$ ، $A_2$ ،... و $A_n$ من $\mathscr{A}$ ، فإنَّ ذلك لا يعني أن تكون العلاقة:

$$P(A_{i_1}A_{i_2}\ldots A_{i_k})=P(A_{i_1})\cdot P(A_{i_2})\cdot\ldots\cdot P(A_{i_k})$$

مُحقَّقةً من أجل جميع قيم الأدلة $i_j \ni \mathbb{N}$ المشار إليها في التعريف السابق.

### (٤,٣,٣) خصائص الحوادث المستقلّة Characteristics of Independent Events

فيما يلي نقدّم مبرهنتين توضّحان لنا بعض الخصائص للحوادث المستقلّ بعضها عن بعض.

**(٤,٣,٣,١) مبرهنة**

ليكن $[\Omega, \mathcal{A}, P]$ فضاءً احتمالياً مُعطى، وليكن $A$ و $B$ حادثين مستقلّين بعضهما عن بعض، فعندئذ:

١- يكون الحادث $A$ مستقلًّا عن $\bar{B}$، وكذلك الحادث $\bar{A}$ مستقلًّا عن $B$ أيضاً.

٢- يكون الحادثان $\bar{A}$ و $\bar{B}$ مستقلّين عن بعضهما.

٣- بفرض أنَّ $A \supseteq A^*$، فعندئذ $A^*$ و $B$ يكونان مستقلّين عن بعضهما **إذا وفقط إذا** كان الحادثان $A \setminus A^*$ و $B$ مستقلّين عن بعضهما.

**البرهان:** من أجل الطلب:

١- سنثبت استقلال $A$ و $\bar{B}$ فقط وذلك لأنَّ إثبات استقلال $B$ و $\bar{A}$ يتم بشكل مماثل، فنجد من كون $A \supseteq AB$ ومن الفرض ما يلي:

$$\begin{aligned} P(A\bar{B}) &= P(A \setminus AB) = P(A) - P(AB) \\ &= P(A) - P(A)\cdot P(B) = P(A)\cdot[1 - P(B)] = P(A)\cdot P(\bar{B}) \end{aligned}$$

وهذا يعني أنَّ $A$ و $\bar{B}$ مستقلان عن بعضهما.

٢- من الفرض وخواص المتمِّمات والعلاقة (٤,٩) يمكننا أن نكتب ما يلي:

$$\begin{aligned} P(\bar{A}\bar{B}) &= P\left(\overline{A \cup B}\right) = 1 - P(A \cup B) = 1 - \left[P(A) + P(B) - P(AB)\right] \\ &= 1 - \left[P(A) + P(B) - P(A)\cdot P(B)\right] \\ &= [1 - P(A)]\cdot[1 - P(B)] = P(\bar{A})\cdot P(\bar{B}) \end{aligned}$$

وهذا يعني أنَّ $\bar{A}$ و $\bar{B}$ مستقلٌ كل منهما عن الآخر.

**٣- أ- من أجل لزوم الشرط:**

لدينا الحادثان $A^*$ و $B$ مستقلان بعضهما عن بعض، فعندئذ يكون:

$$\begin{aligned} P(B)\cdot P(A^*) &= P(A^*B) = P(AB \setminus [(A \setminus A^*) \cap B]) \\ &= P(AB) - P((A \setminus A^*) \cap B) = P(A)\cdot P(B) - P((A \setminus A^*) \cap B) \end{aligned}$$

ومن ثَمَّ ينتج لدينا:

$$P((A \setminus A^*) \cap B) = P(B)\cdot[P(A) - P(A^*)] = P(B)\cdot P(A \setminus A^*)$$

وهذا يعني أنَّ $A \setminus A^*$ و $B$ مستقلٌ كل منهما عن الآخر.

**ب) من أجل كفاية الشرط:**

لدينا الحادثان $A \setminus A^*$ و $B$ مستقلان بعضهما عن بعض، ولنثبت استقلال $A^*$ عن $B$ وذلك على النحو الآتي:

باستخدام خواص الدالّة الاحتمالية $P$، ومن استقلال الحادثين $A$ و $B$ يمكننا أن نكتب الآتي:

$$\begin{aligned} P(B)\cdot P(A \setminus A^*) &= P((A \setminus A^*) \cap B) = P(AB \setminus A^*B) \\ &= P(AB) - P(A^*B) = P(A)\cdot P(B) - P(A^*B) \end{aligned}$$

التي نجد منها أنَّ:

$$P(A^*B) = P(B)\cdot[P(A) - P(A \setminus A^*)] = P(B)\cdot P(A^*)$$

وهذا يعني أنَّ $A^*$ و $B$ مستقلٌ كل منهما عن الآخر.

**(٤,٣,٣,٢) مبرهنة**

ليكن $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً مُعطى، ولتكن $(A_n)_{n\in\mathbb{N}}$ متتالية حوادث متنافية مثنى مثنى من $\mathscr{A}$ ، فعندئذ إذا كان من أجل حادث $B$ من $\mathscr{A}$ لدينا الحوادث $A_n$ و $B$ مستقلّة من أجل أيّة قيمة $n \in \mathbb{N}$ فإنَّه سيكون $B$ مستقل عن $\bigcup_{n=1}^{\infty} A_n$ أيضاً.

**البرهان:** من أجل أي $n \in \mathbb{N}$ لدينا الحادث $A_n$ مستقل عن $B$، ومن ثمَّ باستخدام خواص الدالَّة الاحتمالية $P$ والفرض المقدَّم يكون لدينا ما يلي:

$$\begin{aligned} P\left(B \cap \bigcup_{n=1}^{\infty} A_n\right) &= P\left(\sum_{n=1}^{\infty} (B A_n)\right) = \sum_{n=1}^{\infty} P(B A_n) = \sum_{n=1}^{\infty} P(B)\cdot P(A_n) \\ &= P(B)\sum_{n=1}^{\infty} P(A_n) = P(B)\cdot P\left(\bigcup_{n=1}^{\infty} A_n\right) \end{aligned}$$

وهذا يعني أنَّ $\bigcup_{n=1}^{\infty} A_n$ و $B$ مستقلٌ كل منهما عن الآخر.

**(٤,٣,٣,٣) ملاحظة**

إنَّ المبرهنة السابقة غير قابلة للعكس بمعنى أنَّه إذا كان $B$ مستقل عن $\bigcup_{n=1}^{\infty} A_n$ فليس بالضرورة أنَّ تكون الحوادث $A_n$ و $B$ مستقلّة من أجل أيّة قيمة $n \in \mathbb{N}$، فلو أخذنا على سبيل المثال $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً مُعطى على النحو الآتي:

$$\Omega = \{\omega_1, \omega_2, \omega_3\} \quad \& \quad \mathscr{A} = 2^{\Omega} \quad \& \quad P(\{\omega_i\}) = \frac{1}{3} \quad ; \forall i \in \mathbb{N}_3$$

ولنأخذ الحوادث البسيطة $A_1 = \{\omega_1\}$ ، $A_2 = \{\omega_2\}$ و $A_3 = \{\omega_3\}$ ، فنجد أنَّ هذه الحوادث متنافية مثنى مع $\Omega = \sum_{i=1}^{3} A_i$ ، ولنأخذ متتالية الحوادث $(A_n)_{n\in\mathbb{N}}$ من $\mathscr{A}$ مع $A_i = \varnothing$ من أجل $i \geq 4$ ، فيكون لدينا $\Omega = \sum_{n=1}^{\infty} A_n$ ، وبما أنَّ الحادث الأكيد $\Omega$ مستقل عن أي حادث $B \in \mathscr{A}$ ، فإنَّ ذلك يعني أنَّ الحادث $\sum_{n=1}^{\infty} A_n$ مستقل عن $B$، فلو أخذنا على سبيل المثال $B = \{\omega_1, \omega_2\}$ لوجدنا الآتي:

$$\frac{1}{3} = P(A_1 B) \neq P(A_1)\cdot P(B) = \frac{2}{9}$$

**(٤,٣,٣,٤) أمثلة**

١ – لدينا نظام اتصالات مكوّن من ثلاث محطّات $S_1$ ، $S_2$ و $S_3$ تعمل كلّ منها بشكل مستقل عن الأخريات، فإذا كان احتمال تعطّل هذه المحطّات خلال السنوات الخمس القادمة على الترتيب هو 0.01، 0.02 و 0.03، فعندئذ:

أ – ما هو احتمال عمل المحطّات الثلاث خلال السنوات الخمس القادمة ؟

ب – ما هو احتمال عمل المحطة $S_2$ وتعطّل $S_1$ و $S_3$ خلال السنوات الخمس القادمة ؟

ج – ما هو احتمال تعطّل محطّتين على الأكثر من المحطّات الثلاث خلال السنوات الخمس القادمة؟

**الحـل:** للإجابة على هذه الأسئلة سنفترض أنَّ $A_1$ ، $A_2$ و $A_3$ هو حادث تعطّل المحطة $S_1$ ، $S_2$ و $S_3$ خلال السنوات الخمس

القادمة على الترتيب، فعندئذ بحسب الفرض ستكون الحوادث $A_1$ ، $A_2$ و $A_3$ مستقلّة عشوائياً، ومنه يكون لدينا من أجل الطلب:

أ - بفرض أنَّ $A$ هو الحادث الذي يُعبِّر عن عمل المحطّات الثلاث خلال السنوات الخمس القادمة، فإنَّه سيكون لدينا $A = \bar{A}_1 \bar{A}_2 \bar{A}_3$ ، ومن ثمَّ بسبب استقلال الحوادث $A_1$ ، $A_2$ و $A_3$ ومن المبرهنة (١ ,٣ ,٣ ,٤) يمكننا أن نكتب الآتي:

$$P(A) = P(\bar{A}_1 \bar{A}_2 \bar{A}_3) = P(\bar{A}_1)\cdot P(\bar{A}_2)\cdot P(\bar{A}_3) = 0.9411$$

ب- بفرض أنَّ $B$ هو الحادث الذي يُعبِّر عن عمل المحطّة $S_2$ وتعطّل $S_1$ و $S_3$ خلال السنوات الخمس القادمة هو $B = A_1 \bar{A}_2 A_3$ ، ومن ثمَّ بسبب استقلال الحوادث $A_1$ ، $A_2$ و $A_3$ ومن المبرهنة (١ ,٣ ,٣ ,٤) يكون لدينا ما يلي:

$$P(B) = P(A_1 \bar{A}_2 A_3) = P(A_1)\cdot P(\bar{A}_2)\cdot P(A_3) = 0.0003$$

ج- بفرض أنَّ $C$ هو الحادث الذي يعبر عن تعطّل محطّتين على الأكثر خلال السنوات الخمس القادمة فإنّه يكون لهذا الحادث العرض الآتي:

$$C = \bar{A}_1\bar{A}_2\bar{A}_3 + \bar{A}_1\bar{A}_2A_3 + \bar{A}_1A_2\bar{A}_3 + A_1\bar{A}_2\bar{A}_3 + \bar{A}_1A_2A_3 + A_1\bar{A}_2A_3 + A_1A_2\bar{A}_3$$

ومن ثمَّ بسبب استقلال الحوادث $A_1$ ، $A_2$ و $A_3$ ومن المبرهنة (٤-٣-٣-١) يكون لدينا ما يلي:

$$P(C) = P(\bar{A}_1)\cdot P(\bar{A}_2)\cdot P(\bar{A}_3) + P(\bar{A}_1)\cdot P(\bar{A}_2)\cdot P(A_3) + P(\bar{A}_1)\cdot P(A_2)\cdot P(\bar{A}_3) + P(A_1)\cdot P(\bar{A}_2)\cdot P(\bar{A}_3)$$
$$+ P(\bar{A}_1)\cdot P(A_2)\cdot P(A_3) + P(A_1)\cdot P(\bar{A}_2)\cdot P(A_3) + P(A_1)\cdot P(A_2)\cdot P(\bar{A}_3) = 0.999994$$

وبذلك اكتمل الحل.

**هذا ما تيسَّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل**

# تمارين الفصل الرابع

١- أثبت صحة العلاقات الآتية:

**a)** $\sum_{k=0}^{n} \binom{n}{k} = 2^n$

**b)** $\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$

**c)** $\sum_{k=0}^{n} \binom{n}{k}(-1)^k = \sum_{k=0}^{n} \binom{n}{k}(-1)^{n-k} = 0$

٢- ليكن $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً مُعطى، وبفرض أنَّ $A$ و $B$ حادثين متنافيين ومستقلّين من $\mathscr{A}$ مع $P(A) = 0.35$، فما هو احتمال الحادث $B$؟

٣- ليكن $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً مُعطى، ولنأخذ $A$ و $B$ حادثين من $\mathscr{A}$ ومحقِّقين للعلاقات الآتية:

$$P(A) = 0.45 \,,\; P(A \setminus B) = 0.25 \quad \& \; P(B \setminus A) = 0.25$$

والمطلوب ما يلي:

أ- حساب الاحتمالات الآتية:

**a)** $P(\bar{A})$ & **b)** $P(B)$ & **c)** $P(A \cap B)$

**d)** $P(A \cup B)$ & **e)** $P(\bar{A} \cup \bar{B})$ & **f)** $P(B \mid A)$

ب- تحقَّق إن كان الحادثان $A$ و B مستقلّين أم لا.

٤- ليكن $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً مُعطى، ولنأخذ $A$ و $B$ حادثين من $\mathscr{A}$ ومحقِّقين للعلاقات الآتية:

$$P(A \cup B) = 0.85 \,,\; P(A \setminus B) = 0.20 \quad \& \; P(B \setminus A) = 0.35$$

والمطلوب ما يلي:

أ- حساب الاحتمالات الآتية:

**a)** $P(A \cap B)$ & **b)** $P(A)$ & **c)** $P(B)$

**d)** $P(A \mid B)$ & **e)** $P(\bar{A} \cap \bar{B})$ & **f)** $P(B \mid A)$

ب- تحقَّق إن كان الحادثان $A$ و B مستقلّين أم لا.

٥- أثبت أنَّ تقاطع أيّة أسرة غير خالية من الجبور فوق $\Omega$ هي من جديد جبر على $\Omega$.

٦- لتكن لدينا $\{\mathscr{A}_i : i \in I\}$ أسرة عدودة من الجبور فوق $\Omega$، فعندئذ أثبت أنَّه يكون $\bigcup_{i \in I} \mathscr{A}_i$ جبراً فوق $\Omega$ إذا كان من أجل أي جبرين $\mathscr{A}_\ell$ و $\mathscr{A}_m$ يوجد جبر $\mathscr{A}_k$ بحيث يكون $\mathscr{A}_k \supseteq \mathscr{A}_\ell \cup \mathscr{A}_m$.

٧- لنفترض أنه لدينا بطاقات لعب مكوَّن من 52 بطاقة، وأنَّنا قمنا بخلط هذه البطاقات جيداً، ومن ثمَّ نسحب أربع بطاقات دفعة واحدة من هذه البطاقات، فما هو احتمال:

أ- أن تكون البطاقات من أنواع مختلفة؟

ب- أن تكون البطاقات من نوع واحد؟

ج- أن يكون لدينا بطاقتان من كل لون؟

د- أن تكون البطاقات من ألوان مختلفة؟

٨- لنفترض أنه لدينا بطاقات لعب مكوَّنة من 52 بطاقة، وأنَّنا قمنا بخلط هذه البطاقات جيداً، ومن ثمَّ نسحب أربع بطاقات واحدة تلو الأخرى من هذه البطاقات، فما هو احتمال:

أ- أن تكون البطاقات من أنواع مختلفة؟

ب- أن تكون البطاقات من نوع واحد؟

ج- أن يكون لدينا البطاقتان الأوليان حمراوين؟

د- أن يكون لدينا بطاقة حمراء ثمَّ سوداء ثمَّ حمراء وأخيراً سوداء؟

٩- لنأخذ تجربة إلقاء قطعة نقود متوازنة لأربع مرّات متتالية، فعندئذاجب عما يلي:

أ- عيّن الفضاء الاحتمالي لهذه التجربة العشوائية.

ب- احسب احتمال الحصول على ثلاث صور فقط.

ج- احسب احتمال الحصول على صورتين متتاليتين.

د- هل حادث الحصول على صورتين متتاليتين مستقل عن حادث الحصول على شعارين متتاليتين؟

هـ- ما هو احتمال الحصول على صورتين متتاليتين علماً أنَّنا قد حصلنا على صورتين؟

١٠- لنأخذ تجربة إلقاء قطعة نقود متوازنة لأربع مرات متتالية، ولتكن $A$، $B$ و$C$ هي حوادث الحصول على صورتين متتاليتين في الرمية الأولى، والثانية، والثالثة على الترتيب، فعندئذبيِّن أن كانت هذه الحوادث مستقلّة عشوائياً أم لا؟

١١- لنأخذ تجربة إلقاء حجر نرد متوازن لمرتين متتاليتين، فعندئذ:

أ- عيّن الفضاء الاحتمالي لهذه التجربة العشوائية.

ب- احسب احتمال أن يكون مجموع الرقمين الظاهرين للأعلى أصغر من 6.

ج- احسب احتمال أن يكون الرقمين الظاهرين للأعلى غير متساويين.

د- ما هو احتمال الحصول على رقمين متتاليين؟

هـ- ما هو احتمال أن يكون مجموع الرقمين الظاهرين للأعلى يساوي 10 علماً أنَّنا قد حصلنا على مجموع أكبر من 7؟

و- هل حادث الحصول على الرقم 5 في الرمية الأولى مستقل عن حادث الحصول على الرقم 5 في الرمية الثاني؟

ز- لو أخذنا تجربة إلقاء حجري نرد متوازنين وغير متمايزين لمرّة واحدة فقط، فهل تتغيَّر نتائج الطلبات السابقة، ولماذا؟

١٢- لنأخذ تجربة إلقاء حجري نرد متوازنين و**متمايزين** لمرّة واحدة فقط، فعندئذ أجب عما يلي:

أ- احسب احتمال أن يكون مجموع الرقمين الظاهرين للأعلى أكبر من 6.

ب- احسب احتمال أن يكون الرقمين الظاهرين للأعلى متساويين.

ج- ما هو احتمال الحصول على رقمين غير متتاليين؟

د- ما هو احتمال أن يكون مجموع الرقمين الظاهرين للأعلى أصفر من 8 علماً أنّنا قد حصلنا على مجموع أصغر من 10؟

هـ- هل حادث الحصول على الرقم 3 في الرمية الأولى مستقل عن حادث الحصول على الرقم 2 في الرمية الثاني؟

١٣- في صندوق يوجد 20 كرة متماثلة منها 13 كرة حمراء و5 كرات بيضاء والباقي زرقاء. قمنا بسحب عشوائي لثلاث كرات في آن واحد من هذا الصندوق دون رؤيتها، والمطلوب:

أ- ما هو احتمال أن تكون الكرات الثلاث من ألوان مختلفة؟

ب- ما هو احتمال أن تكون الكرات من لون واحد؟

ج- ما هو احتمال أن يكون لكرتين اللون نفسه والثالثة من لون آخر؟

١٤- في فصل دراسي يوجد 35 طالباً منهم تسعة طلاب متميّزين، وأردنا أن نُشكِّل فريقًا مُكوَّنًا من خمسة طلاب لمسابقة تقام بين مدارس المدينة، فإذا أخذنا خمسة طلاب دفعة واحدة من هذا الفصل وبشكلٍ عشوائي، فعندئذ:

أ- ما هو احتمال أن يكون في هذا الفريق ثلاثة طلاب متميّزين؟

ب- ما هو احتمال أن يكون جميع عناصر اللجنة من الطلاب المتميّزين؟

ج- ما هو احتمال أن يكون في هذا الفريق طالبان متميّزان على الأكثر؟

١٥- لدينا في جهاز $m$ دارة (دائرة كهربائية) تولِّد عشوائياً إشارات منطقية متماثلة وبالاحتمال نفسه، فلو افترضنا أنه لدينا $n$ إشارة ( مع $1 < n < m$ ) فما هو احتمال أن تكون إشارتان على الأقل من هذه الإشارات قد ولِّدت من دارة واحدة؟

١٦- لنفترض أنّه لدينا صندوق $M$ يحوي كرة متماثلة تماماً (أي لها جميعا النصيب نفسه في السحب) ومرقّمة من 1 وحتى $M$ من بينها يوجد $n$ كرة مرقّمة بأعداد إضافية من 1 وحتى $n$ علماً أنَّ $2n \leq M$ ، ولنفترض أن شخصاً ما قام بسحب $n$ كرة دفعة واحدة من الكرات التي في الصندوق، فعندئذ:

أ- ما هو احتمال الحصول على كرة واحدة على الأقل دُوِّن عليها رقم إضافي من بين الكرات ال $n$ التي قام بسحبها؟

ب- إذا كان $M = n^2$ ، فما هو احتمال أن يكون قد حصل على كرة واحدة على الأقل دُوِّن عليها رقم إضافي من الكرات ال $n$ التي قام بسحبها؟

ج- إذا تركنا قيمة $n$ في الطلب السّابق تسعى إلى اللانهاية، فما هي القيمة التي سيأخذها الاحتمال في الطلب السّابق؟

١٧- اتفق شخصان $X$ و$Y$ على الالتقاء في مكان محدَّد وخلال فترة زمنية معيَّنة $[0, t]$ (حيث t مقدرة بالدقيقة)، وبحيث إنَّه إذا وصل الشخص الأول $X$ إلى المكان المحدّد ينتظر $k$ دقيقة، وينصرف في حال عدم وصول الشخص الثاني $Y$، وأمّا إذا وصل $Y$ للمكان المحدّد فإنَّه ينتظر $\ell$ دقيقة، وينصرف في حال عدم وصول $X$، مع العلم أنَّ قيمة $0 < k \leq t$ ، فعندئذ المطلوب:

أ- ما هو احتمال وصول $X$ قبل $Y$ ؟

ب- ما هو احتمال التقاء $X$ مع $Y$ ؟

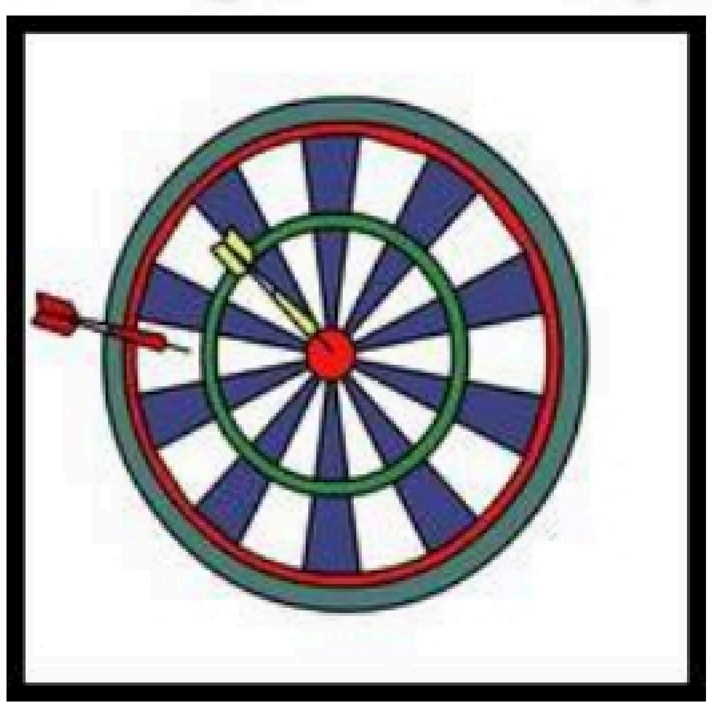
ج- بفرض أنه تمّ الالتقاء فما هو احتمال وصول $X$ قبل $Y$؟

١٨- لنفترض أنه لدينا جسيم يتحرك بسرعة منتظمة ومستقيمة وبشكل موازٍ لمحور الزمن **oT** وأنه خلال كل واحدة زمن (ولتكن الثانية مثلاً) يقوم هذا الجسيم بقفزة نحو الأعلى أو نحو الأسفل بمقدار وحدة طول (ولتكن سنتمترًا مثلا) وباحتمال قدره $0 < p < 1$ و $q := 1 - p$ على الترتيب (إنَّ حركة هذا الجسيم تُعرف باسم الخطى العشوائية Random Walk). فإذا أنجز هذا الجسيم $n$ قفزة، فعندئذ عيِّن الفضاء الاحتمالي لهذه التجربة العشوائية؟

١٩- لنفترض أنه لدينا بطاقات لعب Playing Carts مكوَّنة من 52 بطاقة. نقوم بخلط البطاقات جيداً ومن ثمَّ نسحب خمس بطاقات، فما هو احتمال أن نحصل على العرض الجانبي للبطاقات المسحوبة إذا كنّا:

أ- قد سحبنا البطاقات دفعة واحدة ؟

ب- قد سحبنا البطاقات واحدة تلو الأخرى؟



٢٠- يقوم شخص برمي سهم على لوحة مربَّعة طول ضلعها 56 سم، وفي أوسطها مجموعة من الدائر مركزها مبدأ الإحداثيات (كما هو موضَّح جانباً)، فإذا علمت أنَّ لجميع نقاط لوحة النصيب نفسه في الإصابة، وأنَّه سيصيب نقطة ما من هذه اللوحة بكل تأكيد، وأنَّ أنصاف أقطار هذه الدوائر (مقدَّرةً بـ سم) هي: 1، 10، 11، 20، 21 و23، فعندئذ:

أ- ما هو احتمال أن يصيب القرص الدائري الصغير الذي في المركز؟

ب- ما هو احتمال أن يصيب الشريط الدائري الأوسط (قطره الخارجي 11سم)؟

ج- ما هو احتمال أن يصيب الشريط الدائري الذي قطره الخارجي 20 سم؟

د- ما هو احتمال ألّا يصيب أية نقطة من القرص الدائري؟

# الفصل الخامس

## المتغيرات العشوائية وتوزيعاتها الاحتمالية

## RANDOM VARIABLE AND THEIR PROBABILITY DISTRIBUTIONS

لقد قدّمنا في الفصل السّابق مفهوم الفضاء الاحتمالي $[\Omega, \mathcal{A}, P]$ وكيفية تعيينه، ومن ثمَّ حساب احتمالات لحوادث تتعلّق بتجربة عشوائية، وبعد ذلك انتقلنا إلى الاحتمالات الشرطية واستقلال الحوادث. لكن الوقوف عند هذه الدراسات لا تفي بمتطلبات الكثير من الظواهر الطبيعية والتقنية الموجودة في الواقع حيث تظهر لنا حوادث تنتج عن دوال معرَّفة على مجموعة نتائج تجربة عشوائية. إنَّ الكائن الرياضياتي الذي يعالج هذه المسألة يُعرف بالعنصر العشوائي (وكحالة خاصّة منه المتغيّر العشوائي) الذي سيكون موضع اهتمامنا في هذا الفصل. إنَّ مفهوم العنصر العشوائي يُعدّ واحداً من أهمّ المفاهيم المتعلّقة بنظرية الاحتمالات، ولكن وكما ذكرنا قبل قليل، فإنَّ جلَّ دراستنا في هذا الفصل والفصول القادمة أيضاً ستكون معتمدة على مفهوم المتغيّر العشوائي، ولذلك سنبدأ في توضيح وتعريف المتغيّر العشوائي ومن ثمَّ نعرِّج بشكل سريع وموجز على مفهوم العنصر العشوائي.

### (٥,١) المتغيرات العشوائية

### Random Variables

في الحقيقة يقوم المتغيِّر العشوائي بدور مركزي في نظرية الاحتمالات وتطبيقاتها، وذلك لأنَّ عدداً كبيراً جداً من المشاهدات والملحوظات المعمليّة والطبيعية على حدٍّ سواءٍ يمكن وصفها من خلال متغيِّرات عشوائية أو لأسر (قد تكون عدودة- أي قابلة للعدّ- أو غير عدودة) من المتغيِّرات العشوائيّة، ولهذا يأتي هذا المفهوم في مقدِّمة الدراسات للنظرية الحديثة للاحتمالات. لكن قبل تقديم التعريف لهذا المفهوم سنقدِّم توضيحاً لكيفية نشوء مفهوم المتغيّر العشوائي من خلال مثال عملي ومألوف والمتمثِّل بتجربة قذف قطعة نقود لثلاث مرات متتالية.

لنأخذ تجربة قذف قطعة نقود متوازنة لثلاث مرّات متتالية، فعندئذ الفضاء الاحتمالي $[\Omega, \mathcal{A}, P]$ لهذه التجربة العشوائيّة له المكوِّنات الآتية:

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\} \qquad \& \qquad \mathcal{A} = 2^{\Omega}$$

$$\text{And } P(A) = \sum_{\omega \in A} P(\{\omega\}) \qquad ; \forall A \in \mathcal{A}$$

فلو افترضنا أنَّ $A$ هو الحادث الذي يُعبِّر عن حصولنا على صورتين في الرميات الثلاث، أي إنَّ:

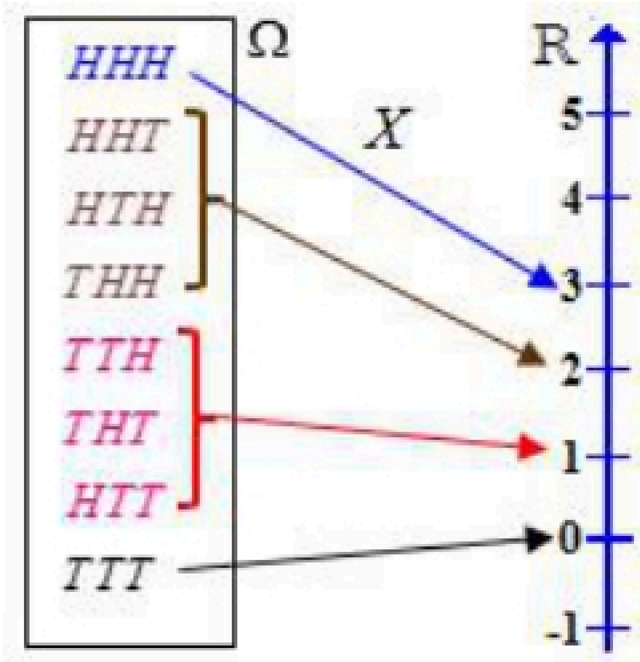
$$A = \{HHT, HTH, THH\}$$

فعندئذ نلاحظ أنّه لا أهميّة لتوضعات الصور خلال هذه الرميات، وهذا يعني أنَّه في مثالنا هذا لا نهتم في أيّة قذفة من هذه القذفات سنحصل على الصورة، وإنّما الذي يهمنا حقاً هو إن كنّا سنحصل في نهاية القذفات الثلاث على صورتين فقط؛ ولذلك دراسة كل حادث

ابتدائي $\omega$ من $\Omega$ لم تَعد مهمّة بالنسبة إلينا في هذه الحالة. الآن لو أمعنا النظر في عبارة الحصول على صورتين فقط، فإنَّه قد يتبادر إلى ذهننا مباشرة فكرة استخدام الأعداد للتعبير عن الحادث $A$ وذلك من خلال عمليّة إرفاق كل حادث ابتدائي $\omega$ من $\Omega$ بعدد يُعبّر عن عدد الصور في هذا الحادث الابتدائي، فيكون لدينا في هذه الحالة المجموعة $\{2\}$ توافق الحادث $A$ (انظر الشكل الآتي).

في الحقيقة إنَّ عملية الإرفاق هذه ما هي إلا تطبيق Map $X$ مُعرَّف على $\Omega$ ويأخذ قيمه في مجموعة الأعداد الحقيقية $\mathbb{R}$ كما يوضِّحه لنا الشكل (١ , ٥) الآتي، وكما هو واضح فإنَّ هـذا التطبيق وحيد القيمة ومعيَّن بشكل وحيد وفقاً للطريقة التي قُدِّم بها، وهذا يعني أنَّ التطبيق $X$ هو في الواقع عبارة عن دالَّة حقيقية مُعرَّفة على $\Omega$.

الآن لو رمزنا لمجموعة قيم التطبيق $X$ بـ $\Omega^*$ فإنَّه سيكون لدينا $\Omega^* = \{0,1,2,3\}$ ومنه نلاحظ أنَّ كل عنصر من $\Omega^*$ هو من جديد حادث ابتدائي لأن ظهور القيم 0، 1، 2 و3 في $\Omega^*$ عشوائي، وعشوائيتها تنتج من عشوائية الحوادث الابتدائية في $\Omega$.



الشكل (١ , ٥)

**ممّا سبق يمكننا القول:**

إنَّ التطبيق $X$ قام بنقل التجربة العشوائيّة التي مجموعة نتائجها $\Omega$ إلى تجربة عشوائية أخرى مجموعة نتائجها $\Omega^*$ التي عناصرها من $\mathbb{R}$ ، ولكن حتى تكون هذه التجربة العشوائيّة الجديدة (التي مجموعة نتائجها $\Omega^*$ ) مُمثِّلة للتجربة العشوائيّة الأصل (مكافئة لها) يجب أن يكون التطبيق $X$ أميناً في (أو محافظاً على) نقل الحوادث أيضاً، فلو كان $\mathcal{A}^*$ هو $\sigma-$ **جبر** فوق $\Omega^*$ (حيث يؤخذ $\mathcal{A}^* = \Omega^* \cap \Re$ ) فإنَّ ذلك يستوجب أن تكون الصورة العكسيّة لأيّ حادث $B$ من $\mathcal{A}^*$ وفقاً للتطبيق $X$ هي من جديد حادث في $\Omega$، فعلى سبيل المثال لو أخذنا $B = \{2\}$ من $\mathcal{A}^*$ في مثالنا المذكور آنفًا فإنَّه يجب أنْ يكون الحادث الآتي:

$$X^{-1}(\{B\}) = X^{-1}(\{2\}) = \{\, \omega \in \Omega \,;\, X(\omega) = 2 \,\} = \{HHT, HTH, THH\}$$

هو حادث من $\mathcal{A}$ ، وهذا مُحقَّق لأنَّ الحادث $\{HHT, HTH, THH\}$ ينتمي إلى $\mathcal{A}$ كما نعلم، ولكن تَحقُّق هذه العلاقة فقط لا يكفي، إذ إنَّنا ذكرنا بأنَّه يجب تحقُّق هذه العلاقة من أجل أي حادث $B$ من $\mathcal{A}^*$ . إنَّ التطبيقات التي تُحقِّق مثل هذه الخاصية تُدعى متغيِّرات عشوائية، وأصل التسمّية تعود إلى كون القيم التي يأخذها التطبيق $X$ (وهو دالَّة في مثالنا هذا) تتغيّر بشكل عشوائي. إنَّ التعريف الرياضياتي الدقيق لهذا المفهوم تقدِّمه لنا الفقرة الآتية.

### (١ , ١ , ٥) تعريف (المتغيِّر العشوائي)

ليكن $[\Omega, \mathcal{A}, P]$ فضاءً احتمالياً مُعطى، و $[\mathbb{R}, \Re]$ فضاءً مقيساً، ولنأخذ $X : \Omega \longrightarrow \mathbb{R}$ تطبيقاً ما، فإذا كان من أجل أيّة مجموعة بوريلية $B \ni \Re$ لدينا:

$$X^{-1}(B) = \{\, \omega \in \Omega \,;\, X(\omega) \in B \,\} \in \mathcal{A} \qquad [5,1]$$

فعندئذ يُقال عن $X$ إنَّه متغيِّر عشوائي فوق الفضاء الاحتمالي $[\Omega, \mathcal{A}, P]$ .

إنَّ العلاقة السّابقة تعني (في نظرية القياس) أنَّ التطبيق $X$ قيوساً (أو مقيساً أو قابلاً للقياس Measurable) بالنسبة إلى $\mathcal{A}$ و $\Re$ ، وهذا يعني أنَّ المتغيرات العشوائيّة في حقيقتها هي **تطبيقات قيوسة** Measurable Maps.

في الجوانب التطبيقية يُعدّ تحقيق العلاقة [5,1] ليس سهلاً في مُعظم الحالات، وفي حال كان ذلك ممكنًا فهو شاقّ جداً، ولذلك قُدّمت اختبارات لتحديد إن كان تطبيق $X$ فوق فضاء احتمالي $[\Omega, \mathcal{A}, P]$ متغيِّرًا عشوائيًا أم لا.

إنَّ المبرهنة الآتية تقدِّم لنا أحد هذه الاختبارات، علماً أنَّنا سنقبلها دون برهان لأنَّ برهانها يقع في نطاق نظرية القياس.

**(٥,١,١,١) مبرهنة**

ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً مُعطى، ولنأخذ $X:\Omega\to\mathbb{R}$ تطبيقاً حقيقياً معرّفاً على $\Omega$ بحيث $\omega\mapsto X(\omega)$، فإذا كان:

$$\{\omega\in\Omega\,;X(\omega)<x\}\in\mathscr{A}\qquad;\forall x\in\mathbb{R}\qquad [5,2\text{-a}]$$

فعندئذ يكون $X$ متغيِّراً عشوائياً فوق $[\Omega,\mathscr{A},P]$.

**(٥,١,١,٢) ملاحظات**

١- فيما بعد سنعتمد على الاختبار الوارد في المبرهنة السّابقة لتبيان إن كان تطبيق $X$ فوق فضاء $[\Omega,\mathscr{A},P]$ متغيِّرا عشوائيًّا أم لا، وقد وقع اختيارنا على هذا الاختبار بسبب حاجتنا للحادث $\{\omega\in\Omega\,;X(\omega)<x\}$ في تعريف دالَّة التوزيع الاحتمالية لمتغيِّر عشـوائي $X$ لاحقاً.

٢- بما أنَّ $\{\omega\in\Omega\,;X(\omega)<x\}=X^{-1}((-\infty\,,x))$ فإنَّ الشرط السّابق [5,2-a] يكافئ الشرط الآتي أيضاً:

$$X^{-1}((-\infty\,,x))\in\mathscr{A}\qquad;\forall x\in\mathbb{R}\qquad [5,2\text{-b}]$$

٣- على سبيل الاختصار والتبسيط يُكتب في كثير من الحالات $P(X<x)$ عوضاً عن:

$$P(\{\omega\in\Omega\;\;;X(\omega)<x\})$$

وعندما نكتب الصيغة المفصَّلة إنَّما نقصد بذلك التذكير بها أو لضرورة تُمليها طبيعة العرض للعلاقة التي هي قيد الدراسة.

٤- لكي نتجنب سوء الفهم الذي يمكن أن ينشأ عن مصطلح المتغيِّر العشوائي كدالَّة نودّ أن نشير هنا إلى أنَّ المتغيِّر العشوائي $X$ كدالَّة إنّما يعني أنّه لدى متغيِّرات مستقلَّة $\omega$ من $\Omega$ مُعطاة تكون القيم $X(\omega)$ وحيدة التعيين، وأنَّ قبول هذه الدالَّة كمتغيِّر عشوائي يتوقف على اختيار المتغيِّرات المستقلَّة $\omega$ من $\Omega$، وعلى اختيار الـ $\sigma$- **جبر** $\mathscr{A}$، والأمثلة الآتية توضِّح ذلك.

**(٥,١,١,٣) أمثلة**

١- ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً مُعطى، و$X$ تطبيقاً حقيقياً معرّفاً على $\Omega$ من خلال العلاقة الآتية:

$$X(\omega)=c\qquad;\forall\,\omega\in\Omega$$

علماً أنَّ $c$ عدد حقيقي ثابت (أي إنَّ $X$ هو التطبيق الثابت)، فعندئذ نجد أنَّ هذا التطبيق هو متغيِّر عشـوائي فوق $[\Omega,\mathscr{A},P]$ وذلك لأنَّ:

$$\{\omega\in\Omega\;\;;X(\omega)<x\}=\begin{cases}\varnothing & for\;\;x\le c\\ \Omega & for\;\;x>c\end{cases}$$

حيث نعلم أنَّ كلاً من $\varnothing$ و$\Omega$ هي عناصر من $\mathscr{A}$ دوماً لأنَّ $\mathscr{A}$ هو $\sigma$- **جبر** فوق $\Omega$. إذاً التطبيق $X$ هو متغيِّر عشـوائي فوق $[\Omega,\mathscr{A},P]$، ويُدعى **متغيِّر عشوائي وحيد النقطة** (أو الوحيد القيمة أو الثابت) فوق $[\Omega,\mathscr{A},P]$ أيضاً. لاحظ هنا أنّه حتى لو أخذنا $\mathscr{A}$ في نموذج الفضاء الاحتمالي $[\Omega,\mathscr{A},P]$ هو الجبر الضعيف $\mathscr{A}=\{\varnothing,\Omega\}$، فإنَّ التطبيق $X$ المُعطى في هذا المثال هو متغيِّر عشوائي فوق $[\Omega,\mathscr{A},P]$، ولكن يجب الانتباه إلى أنَّ هذا الطرح ليس صحيحاً في الحالة العامّة، والمثال التالي يوضّح لنا ذلك.

٢- ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً مُعيَّناً على النحو الآتي:

$$\Omega=\{\omega_1\,,\omega_2\}\qquad\&\qquad\mathscr{A}=2^{\Omega}$$

$$P(A)=p\cdot\delta_{\omega_1}(A)+q\cdot\delta_{\omega_2}(A)\qquad;\forall A\in\mathscr{A},\;\;p+q=1\,.$$

أي إنّه لدينا $P(\{\omega_1\})=p$ و $P(\{\omega_2\})=q$ (راجع المبرهنة(٣-٤-١-٤) في الفصل السّابق)، ولنأخذ $X$ تطبيقاً حقيقياً معرّفاً على $\Omega$ كما يلي:

$$X(\omega)=\begin{cases}0 & for \quad \omega=\omega_2\\ 1 & for \quad \omega=\omega_1\end{cases}$$

ولنتبيّن إن كان هذا التطبيق هو متغيِّر عشوائي فوق $[\Omega,\mathscr{A},P]$ أم لا. من أجل ذلك لدينا:

$$\{\omega\in\Omega\,;X(\omega)<x\}=\begin{cases}\varnothing & for \quad x\le 0\\ \{\omega_2\} & for \quad 0<x\le 1\\ \{\omega_1,\omega_2\}=\Omega & for \quad x>1\end{cases}$$

ولكن نعلم أنَّ $\{\omega_2\}$، $\varnothing$ و$\Omega$ هي عناصر من $\mathscr{A}$، ومن ثَمَّ التطبيق $X$ هو متغيِّر عشوائي فوق $[\Omega,\mathscr{A},P]$.

لاحظ أنَّه لو أخذنا $\mathscr{A}=\{\varnothing,\Omega\}$ في نموذج الفضاء الاحتمالي $[\Omega,\mathscr{A},P]$ (وهو $\sigma$- **جبر** فوق $\Omega$ كما نعلم)، فعندئذ التطبيق $X$ ليس متغيِّراً عشوائياً فوق $[\Omega,\mathscr{A},P]$ لأنَّ $\{\omega_2\}$ ليس عنصراً من $\mathscr{A}$ في هذه الحالة. لذلك عندما نريد تبيان إن كان تطبيق $X$ هو متغيِّر عشوائي فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ فإنَّه علينا الانتباه إلى الـ $\sigma$- **جبر** الذي نتعامل معه.

كتطبيق على هذا المثال لنأخذ تجربة قذف قطعة نقود (ليس بالضرورة أنَّ تكون متوازنة) لمرة واحدة فقط، وليكن $X$ تطبيقاً حقيقياً معـرَّفاً على مجموعة نتائج هذه التجربة العشوائية كما يلي:

$$X(\omega)=\begin{cases}0 & for \quad \omega=H\\ 1 & for \quad \omega=T\end{cases}$$

فنجد أنَّ:

$$\{\omega\in\Omega\,;X(\omega)<x\}=\begin{cases}\varnothing & for \quad x\le 0\\ \{H\} & for \quad 0<x\le 1\\ \{H,T\}=\Omega & for \quad x>1\end{cases}$$

وبما أنَّ $\{H\}$، $\varnothing$ و $\Omega$ هي عناصر من $\mathscr{A}=2^{\Omega}$ فإنَّ ذلك يعني أنَّ التطبيق $X$ هو متغيِّر عشوائي فوق الفضاء الاحتمالي لهذه التجربة العشوائيّة.

٣- ليكن $[\Omega,\mathscr{A},P]$ هو الفضاء الاحتمالي لتجربة إلقاء حجر نرد متوازن لمرّتين متتاليتين، أي إنَّ:

$$\Omega=\{\omega=(\omega_1,\omega_2)\;;\omega_1,\omega_2\in\{1,2,\dots,6\}\}\quad\&\quad \mathscr{A}=2^{\Omega}\quad\&\quad P(A)=\sum_{\omega_1,\omega_2=1}^{6}\frac{1}{36}\,\delta_{(\omega_1,\omega_2)}(A)\quad;\forall A\in\mathscr{A}$$

وليكن $X$ تطبيقاً معرَّفاً فوق $\Omega$ من خلال العلاقة الآتية:

$$X(\omega)=\begin{cases}0 & for \quad \omega_1+\omega_2<10\\ 1 & for \quad \omega_1+\omega_2\ge 10\end{cases}$$

فنجد أنَّ هذا التطبيق هو متغيِّر عشوائي فوق $[\Omega,\mathscr{A},P]$ وذلك لأنَّ:

$$\{\omega\in\Omega\,;X(\omega)<x\}=\begin{cases}\varnothing & for \quad x\le 0\\ \{(\omega_1,\omega_2)\in\Omega\,;\omega_1+\omega_2<10\} & for \quad 0<x\le 1\\ \Omega & for \quad x>1\end{cases}$$

حيث نعلم أنَّ $\varnothing$، $\Omega$ من $\mathscr{A}$ دوماً وكذلك $\{(\omega_1,\omega_2)\in\Omega\,;\omega_1+\omega_2<10\}$ هو عنصر من $2^{\Omega}$ أيضاً.

٤- يقف شخص على شرفة بناء شاهق (مرتفع جداً) مَبنيٌّ على حافة نهر ضيق المسار بحيث يبدو للشخص على شكل خط مستقيم. يلقي هذا الشخص بحصى على النهر بحيث يمكنه أن يمسح جزءًا من طول النهر المحدود بنقطتين $a$ و$b$، علماً أنَّ $-\infty < a < b < \infty$ ، فلو افترضنا أنَّ الحصى الملقاة ستصيب نقطة ما من النهر (نقطة من الفترة $[a , b]$) حتماً، وكذلك لجميع نقاط هذه الفترة النصيب نفسه في الإصابة، فعندئذ سيكون الفضاء الاحتمالي $[\Omega, \mathscr{A}, P]$ لهذه التجربة العشوائيّة معيّن من خلال المُعطيات الآتية:

$$\Omega = [a , b] \qquad \& \qquad \mathscr{A} = \Re \cap [a , b] \qquad \& \qquad P(A) = \frac{1}{b-a}\int_A dx \quad ; \forall A \in \mathscr{A}$$

وليكن $X$ تطبيقاً حقيقياً مُعرَّفاً على $\Omega$ كما يلي:

$$X(\omega) = \omega \quad ; \forall \omega \in \Omega$$

فنجد أنَّ هذا التطبيق (التطبيق المطابق على $\Omega$) هو متغيِّر عشوائي فوق $[\Omega, \mathscr{A}, P]$ وذلك لأنَّ:

$$\{ \omega \in \Omega \; ; X(\omega) < x \} = \begin{cases} \varnothing & for \quad x \le a \\ [a , x) & for \quad a < x \le b \\ [a , b] = \Omega & for \quad x > b \end{cases}$$

حيث نعلم أنَّ $\varnothing$ و$\Omega$ من $\mathscr{A}$ دوماً، والفترات $[a , x)$ مع $x \le b$ هي مجموعات بوريلية من $\Re$ واقعة في الفترة $[a , b]$ ، ومن ثَمّ فإنَّ $[a , x)$ عنصر من $\mathscr{A}$ أيضاً.

### (٤ , ١ , ١ , ٥) ملاحظات

١- بالعودة إلى المثال / ١/ من الفقرة السّابقة نلاحظ أنَّ:

$$P(X = c) = P(\{ \omega \in \Omega \; ; X(\omega) = c \}) = P(\Omega) = 1$$

ولذلك يُقال إنَّ $X$ متغيِّر عشوائي ثابت باحتمال يساوي الواحد (أو تقريباً أكيد ) أيضاً، ويُكتَب ذلك على النحو الآتي:

$$X \overset{a.s.}{=} c \quad \textbf{or} \quad X = c \;\; ; a.s. \quad \textbf{or} \quad X = c \quad almost\ all \;\; \omega \in \Omega$$

علماً أنَّ $(a.s.)$ هي اختصار لـ Almost Sure ، ولكن يجب الانتباه هنا إلى أنَّ العكس ليس صحيحاً في الحالة العامّة، بمعنى أنَّه إذا كان لدينا $X \overset{a.s.}{=} c$ فإنَّ ذلك لا يعني أنَّ $X(\omega) = c$ من أجل كل $\omega \in \Omega$ ، وذلك لإنَّه من الممكن أن يوجد حادث $\mathscr{A} \ni A \ne \varnothing$ بحيث $X(\omega) \ne c$ من أجل كل أي $\omega \in \Omega$ ، ولكن في هذه الحالة يجب أن يكون $P(A) = 0$ ، أي إنَّ الحادث $A$ في هذه الحالة هو حادث شبه مستحيل.

٢- إذا كان $X$ و$Y$ متغيِّرين عشوائيين فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ ، وكان من أجلهما لدينا:

$$P(X = Y) = P(\{ \omega \in \Omega \; ; X(\omega) = Y(\omega) \}) = 1 \qquad [5,3]$$

فعندئذ يقال عن هذين المتغيِّرين العشوائيين إنَّهما متكافئان ويكتَب ذلك بالشكل $X \approx Y$ ، أو يُقال إنَّهما متساويان باحتمال يساوي الواحد، ويكتَب ذلك على النحو الاتي:

$$X \overset{a.s.}{=} Y \qquad \textbf{or} \qquad X = Y \quad almost\ all \;\; \omega \in \Omega \qquad [5,4]$$

وبناء على ذلك فإنَّه يمكن النظر إلى القيم الثابتة من $\mathbb{R}$ على أنَّها متغيِّرات عشوائية ثابتة معرَّفة على $\Omega$ وتأخذ قيمتها نفسها باحتمال يساوي الواحد.

٤- كما رأينا في الملاحظة السّابقة / ٢/ ، فإنَّ المتغيِّرات العشوائيّة تتساوى فيما بينها على حوادث شبه أكيدة (أو على الحادث الأكيد

في بعض الحالات الخاصّة) وذلك بسبب تحقُّق العلاقة [5,3]، ولكن على سبيل التبسيط والاختصار سنكتب $X=Y$ دون أن نذكر عبارة "باحتمال يساوي الواحد" ، حيث يجب أن يُفهم من هذه الكتابة دوماً أنَّهما متساويان باحتمال يساوي الواحد.

لقد ذكرنا سابقاً أنَّ مفهوم المتغيِّر العشوائي ما هو إلاّ حالة خاصّة من مفهوم أكثر شمولية يُعرف باسم العنصر العشوائي Random Element، ويقدمه لنا التعريف الآتي.

**(٥ , ١ , ٢) تعريف (العنصر العشوائي Random Element)**

ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً و $[G,\mathfrak{G}]$ فضاءً مقيساً ذو طبيعة ما، ولنأخذ $X:\Omega \longrightarrow G$ تطبيقاً قيوساً بالنسبة إلى $\mathscr{A}$ و $\mathfrak{G}$ ، فعندئذ يُسمَّى التطبيق $X$ **عنصراً عشوائياً** فوق الفضاء الاحتمالي $[\Omega,\mathscr{A},P]$ ، أو يُقال إنَّ $X$ هو متغيِّر عشوائي فوق الفضاء الاحتمالي $[\Omega,\mathscr{A},P]$ وبقيم في الفضاء المقيس $[G,\mathfrak{G}]$ . أي إنَّ:

$$X \text{ is a random element } \Leftrightarrow X^{-1}(B)\in\mathscr{A}\ \ ;\forall B\in\mathfrak{G} \qquad [5,5]$$

عادةً تُخصَّص كلمة متغيِّر عشوائي (دون ذكر بقيم في الفضاء المقيس $[G,\mathfrak{G}]$ ) للعناصر العشوائيّة التي تأخذ قيمها في $[\mathbb{R},\mathfrak{R}]$ فقط، وفيما عدا ذلك يكون لـ $X$ تسميات أخرى حسب بُنْية الفضاء المقيس $[G,\mathfrak{G}]$ ، والأمثلة الآتية تعرض لنا بعض النماذج الخاصّة من العناصر العشوائيّة.

**(٥ , ١ , ٢ , ١) أمثلة**

ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً و $[G,\mathfrak{G}]$ فضاءً مقيساً، ولنأخذ $X:\Omega\longrightarrow G$ تطبيقاً قيوساً بالنسبة إلى $\mathscr{A}$ و $\mathfrak{G}$ ، فعندئذ:

١ – إذا أخذنا $[G,\mathfrak{G}]=[\mathbb{R},\mathfrak{R}]$ فإنَّ $X$ يُدعى **متغيِّراً عشوائياً** Random Variable (كما نوهنا بذلك قبل قليل).

٢ – إذا أخذنا $[G,\mathfrak{G}]=[\mathbb{R}^n,\mathfrak{R}^n]$ فإنَّ $X$ يُدعى **متَّجهاً عشوائياً** بُعده $n$ ($n$-Dimensional Random Vector)، وسنرمز له بـ $\mathbb{X}_n$ .

٣ – إذا أخذنا $[G,\mathfrak{G}]=[\mathbb{C},\mathfrak{C}]$ فإنَّ $X$ يُدعى **متغيِّراً عشوائياً عقديًا** Complex Random Variable، علماً أنَّ $\mathbb{C}$ هي مجموعة الأعداد العقديّة (أو المركَّبة) و $\mathfrak{C}$ هو أصغر $\sigma$ – **جبر** على $\mathbb{C}$ الذي يحوي المجموعات ذات الشكل الآتي:

$$\{z\in\mathbb{C}\ ;z=x+\mathbf{i}\,y\ \ ;a_1\le x<b_1\ \&\ a_2\le y<b_2\ with\ \ a_1,a_2,b_1,b_2\in\mathbb{R}\}$$

٤ – إذا أخذنا $[G,\mathfrak{G}]=[\mathbf{M},\mathcal{M}]$ علماً أنَّ الفضاء المقيس $[\mathbf{M},\mathcal{M}]$ يُبنى على النحو الآتي:

نأخذ $[\mathbf{X},\rho]$ فضاءً مترياً فصولاً (قابلاً للفصل Separable) وتاماً، وليكن $\mathfrak{X}$ أصغر $\sigma$ – **جبر** على $\mathbf{X}$ . كذلك لنرمز بـ $\mathbf{M}$ لمجموعة كل القياسات $\varphi$ على $[\mathbf{X},\mathfrak{X}]$ بحيث يكون من أجل كل $A\in\mathfrak{X}$ لدينا $\varphi(A)\in\overline{\mathbb{N}}$ ، ومن ثمَّ نأخذ $\mathcal{M}$ هو أصغر $\sigma$ – **جبر** على $\mathbf{M}$ الذي من أجله التطبيقات الآتية قيوسة من أجل كل $A\in\mathfrak{X}$ :

$$\Delta_A:\mathbf{M}\longrightarrow\overline{\mathbb{N}}\ \ ;\varphi\mapsto\Delta_A(\varphi):=\varphi(A)$$

عندئذ يُدعى العنصر العشوائي $X$ **طورياً نقطياً** Point Process.

٥ – إذا أخذنا $[G,\mathfrak{G}]=[\mathbf{N},\mathfrak{N}]$ علماً أنَّ الفضاء المقيس $[\mathbf{N},\mathfrak{N}]$ يُشكَّل على النحو الآتي:

ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً و $[\mathbf{S},\mathfrak{S}]$ فضاءً مقيساً، ولتكن $\mathbf{N}$ مجموعة كل قوانين التوزيع (سنأتي على شرح هذا المفهوم لاحقاً في هذا الفصل) على $\mathfrak{S}$ ، وليكن $\mathfrak{N}$ هو أصغر $\sigma$ – **جبر** فوق $\mathbf{N}$ الذي من أجله تكون التطبيقات $\Delta_B$ الآتية قيوسة من أجل كل $B\in\mathfrak{S}$ :

$$\Delta_B : \mathbf{N} \longrightarrow \mathbb{R} \ ; P \mapsto \Delta_B(P) := P(B)$$

عندئذ يُسمَّى العنصر العشوائي $X$ **نواة عشوائية** Random Kern.

٦- إذا أخذنا $[G,\mathfrak{G}] = [\mathbf{F},\mathfrak{F}]$ علماً أنَّ الفضاء المقيس $[\mathbf{F},\mathfrak{F}]$ يُبنى على النحو الآتي:
ليكن $\mathbf{F}$ صف كل المجموعات الجزئية المغلقة في $\mathbb{R}^n$، ومن ثمَّ نأخذ $\mathfrak{F}$ هو أصغر $\sigma$- **جبر** فوق $\mathbf{F}$ والحاوي على كل المجموعات ذات الشكل الآتي:

$$A_K := \{A \in \mathfrak{F} \ ; A \cap K \neq \varnothing\} \ \ ; \forall \ K \in \mathcal{K}$$

علماً أنَّ $\mathcal{K}$ يرمز لصف كل المجموعات المتراصة Compact Sets في $\mathbb{R}^n$، فعندئذ يُدعى العنصر العشوائي $X$ **مجموعة مغلقة عشوائية** Random Closed Set.

سنقدِّم فيما يلي أحد المفاهيم المهمّة التي تستخدم في تقديم العديد من صيغ الدوال والمتغيِّرات العشوائيّة على حدٍّ سواء (من أجل الاطلاع على خصائصها راجع تعريف الدالَّة المميِّزة لمجموعة في الملحق **A**).

### (٥، ١، ٣) الدالَّة المميِّزة لحادث Indicator Function of an Event

**تعريفها:** من أجل $\Omega$ مجموعة نتائج تجربة عشوائية ما، فإنَّ **الدالَّة المميِّزة لمجموعة** $A \subseteq \Omega$ هي دالَّة حقيقية $\mathbf{I}_A$ معرَّفة على $\Omega$ كما يلي:

$$\mathbf{I}_A(\omega) = \begin{cases} 1 & for \quad \omega \in A \\ 0 & for \quad \omega \notin A \end{cases}$$ [5,6]

**(٥، ١، ٣، ١) مبرهنة**

ليكن $[\Omega,\mathcal{A},P]$ فضاء احتمالي مُعطى، ولنأخذ $A \subseteq \Omega$. عندئذ تكون الدالَّة $\mathbf{I}_A$ متغيِّراً عشوائياً فوق $[\Omega,\mathcal{A},P]$ **إذا وفقط إذا** كانت المجموعة $A$ حادثاً من $\mathcal{A}$.

**البرهان:** بملاحظة أنَّ:

$$\{\omega \in \Omega \ ; \mathbf{I}_A(\omega) < x\} = \begin{cases} \varnothing & for \ \ x \leq 0 \\ \overline{A} & for \ \ 0 < x \leq 1 \\ \Omega & for \ \ x > 1 \end{cases}$$

وبما أنَّ $\mathcal{A}$ هو $\sigma$- **جبر** فوق $\Omega$، فإنَّ $\varnothing$ و $\Omega$ هي عناصر من $\mathcal{A}$ دوماً، ولذلك تكون الدالَّة $\mathbf{I}_A$ متغيِّراً عشوائياً فوق $[\Omega,\mathcal{A},P]$ **إذا وفقط إذا** كانت $\overline{A}$ عنصراً من $\mathcal{A}$، والذي يقتضي أن يكون $A$ عنصراً من $\mathcal{A}$ أيضاً (لأنَّ $\mathcal{A}$ مغلق بالنسبة إلى عملية التتميم)، ومن ثَمّ تكون الدالَّة $\mathbf{I}_A$ متغيِّراً عشوائياً فوق $[\Omega,\mathcal{A},P]$ **إذا وفقط إذا** كانت المجموعة $A$ حادثاً من $\mathcal{A}$.

### (٥، ١، ٤) الدوال في متغيّرات عشوائية

المبرهنات الآتية تعرض لنا بعض الخصائص المهمّة للمتغيِّرات العشوائيّة، حيث تستند في ادعاءاتها إلى خصائص الدوال القيوسة (للاطلاع راجع الملحق **B**).

**(٥، ١، ٤، ١) مبرهنات (تقبل دون برهان)**

١- إذا كانت $g$ دالَّة حقيقية مُعرَّفة على $\mathbb{R}$، وكانت هذه الدالَّة متباينة أو مستمرَّة على $\Omega$، فعندئذ تكون هذه الدالَّة مقيسة بالنسبة إلى $\mathcal{A}$ و $\mathfrak{R}$، ومن ثَمّ بفرض أنَّ $X$ متغيِّر عشوائي فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$، فإنَّ $g(X)$ سيكون متغيِّراً عشوائياً فوق الفضاء الاحتمالي $[\Omega,\mathcal{A},P]$ أيضاً.

٢- إذا كان $Y$ متغيّراً عشوائياً فوق فضاء احتمالي $[\mathbb{R},\Re,S]$، وبفرض أنّ $X$ متغيّر عشوائي فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فإنّ التركيب $\Omega \xrightarrow{X} \mathbb{R} \xrightarrow{Y} \mathbb{R}$ هو من جديد متغيّر عشوائي فوق $[\Omega,\mathscr{A},P]$ (لأنّ تركيب تطبيقين قيوسين هو تطبيق قيوس أيضاً).

٣- لتكن $X_1$، $X_2$، ... و $X_n$ متغيّرات عشوائية فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فإذا كانت $g$ دالّة حقيقية مُعرّفة على $\mathbb{R}^n$ ومقيسة بالنسبة إلى $\Re$ و $\Re^n$، فعندئذ تكون الدالّة $Y=g\left(X_1,X_2,\ldots,X_n\right)$ مقيسة بالنسبة إلى $\mathscr{A}$ و $\Re$، ومن ثمّ يكون $Y$ متغيّراً عشوائياً فوق $[\Omega,\mathscr{A},P]$.

تسمح المبرهنة الآتية بدراسة نماذج مختلفة من المتغيّرات العشوائيّة والتعامل مع العمليّات الجبرية على المتغيّرات العشوائيّة، وكذلك دراسة بعض الحالات الحدّية لمتتاليات من المتغيّرات العشوائيّة.

**(٢، ٤، ١، ٥) مبرهنة (يترك إثبات فقراتها كتمارين للقارئ)**

لتكن $X$، $Y$، $X_1$، $X_2$ و... متغيّرات عشوائية فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فعندئذ ستكون الدالّة $Z$ في الصيغ الآتية هي متغيّر عشوائي فوق $[\Omega,\mathscr{A},P]$ أيضاً:

1) $Z=|X|$

2) $Z=aX+b \quad ;\ a,b\in\mathbb{R}\ \ and\ \ a\neq 0$

3) $Z=X\pm Y$

4) $Z=X\cdot Y$

5) $Z=\dfrac{X}{Y} \quad ;Y(\omega)\neq 0\ \ for\ \ all\ \ \omega\in\Omega$

6) $Z=\max\left\{X_i,X_j\right\} \quad ;i,j\in\mathbb{N},i\neq j$

7) $Z=\min\left\{X_i,X_j\right\} \quad ;i,j\in\mathbb{N},i\neq j$

8) $Z=\sup\{X_n\ ;n\in\mathbb{N}\}$

9) $Z=\inf\{X_n\ ;n\in\mathbb{N}\}$

10) $Z=\overline{\lim_{n\to\infty}}X_n=\lim_{n\to\infty}\sup_{m\geq n}X_m$

11) $Z=\underline{\lim_{n\to\infty}}X_n=\lim_{n\to\infty}\inf_{m\geq n}X_m$

12) $Z=\lim_{n\to\infty}X_n$

## (٥،٢) قانون ودالة توزيع متغير عشوائي

## Distribution Law and Function of a Random Variable

من المفاهيم المهمّة التي ترتبط مع مفهوم المتغيّر العشوائي ارتباطاً وثيقاً ومباشراً ما يُعرف باسم **قانون التوزيع** Distribution Law لمتغيّر عشوائي، وأهميّة هذا القانون تكمن في أنّه يصف التوزيع الاحتمالي لقيم المتغيّر العشوائي.

### (٥، ٢، ١) قانون توزيع متغيّر عشـوائي

لقد أُدخل مفهوم قانون التوزيع لأوّل مرّة من قبل الرياضياتي والفيزيائي الفرنسـي **بواسون**-Siméon Denis Poisson (1781-1840) حيث قَدّمه كدالّة احتمالية وذلك لأنّ القياس لم يكن مطروحاً آنذاك. أمّا التعريف الرياضياتي الدقيق لهذا الكائن فيقدّمه لنا التعريف الآتي.

(٥,٢,١,١) تعريف (قانون توزيـع متغيِّر عشوائي)

ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً مُعطى، و $[\mathbb{R},\Re]$ فضاءً مقيساً، ولنأخذ $X:\Omega\longrightarrow\mathbb{R}$ متغيِّراً عشوائياً، ولنضع:

$$P_X := P\circ X^{-1} \qquad [5,7\text{-a}]$$

فنجد أنَّها دالَّة مجموعات حقيقية معرَّفة على $\Re$، أي إنَّ:

$$P_X(B)=P\circ X^{-1}(B) \quad ; B\in\Re \qquad [5,7\text{-b}]$$

علماً أنَّ الرمز ($\circ$) يشير إلى تركيب التطبيقات، فعندئذ $P_X$ يُسمَّى **قانون توزيع** Distribution Law المتغيِّر العشوائي $X$.

(٥,٢,١,٢) ملاحظات

١- يُقال عن $P_X$ إنَّه القياس المولَّد من المتغيِّر العشـوائي $X$، ويمكن التحقّق بسهولة من أنَّ $P_X$ هو قياس احتمالي على $[\mathbb{R},\Re]$، ومن ثمَّ يكون فضاء القياس $[\mathbb{R},\Re,P_X]$ هو فضاء احتمالي، ويُسمَّى **الفضاء الاحتمالي المولَّد من المتغيّر العشوائي $X$** .

٢- من أجل $A$ حادث ما من $\mathscr{A}$ يكون لقانون توزيع المتغيِّر العشوائي $\mathbf{I}_A$ العرض الآتي:

$$P_{\mathbf{I}_A}(B)=\begin{cases} P(A) & for \quad 1\in B\\ P(\overline{A}) & for \quad 0\in B\\ 1 & for \quad 0,1\in B\\ 0 & O.W.\end{cases}$$

(٥,٢,١,٣) أمثلة

١- بالعودة إلى المثال /١/ من (٥,١,١,٣) نجد أنَّ قانون توزيع المتغيِّر العشوائي $X$ يُعيَّن على النحو الآتي:

من أجل أي حادث $B$ (مجموعة بوريلية) من $\Re$ لدينا ما يلي:

$$P_X(B)=P\left(X^{-1}(B)\right)=\begin{cases} P(\Omega)=1 & for \quad c\in B\\ P(\varnothing)=0 & for \quad c\notin B\end{cases}$$

ومن ثمَّ يكون لقانون توزيع هذا المتغيِّر العشـوائي العرض الآتي:

$$P_X(B)=\begin{cases}1 & for \quad c\in B\\ 0 & for \quad c\notin B\end{cases} \qquad ;\forall B\in\Re$$

أو أن نقوم بعرضه من خلال العلاقة الآتية:

$$P_X(\{x\})=\begin{cases}1 & for \quad x=c\\ 0 & for \quad x\neq c\end{cases} \qquad ;\forall x\in\mathbb{R}$$

٢- بالعودة إلى المثال /٢/ من (٥,١,١,٣) نجد أنَّ قانون توزيع المتغيِّر العشوائي $X$ يُعيَّن على النحو الآتي. من أجل أي حادث $B$ من $\Re$ لدينا:

$$P_X(B)=\begin{cases} q=1-p & for \quad 0\in B\\ p & for \quad 1\in B\\ 1 & for \quad 0 \text{ and } 1\in B\\ 0 & for \quad 0 \text{ and } 1\notin B\end{cases}$$

التي يمكن كتابتها على النحو الآتي أيضاً:

$$P_X\left(\{x\}\right)=\begin{cases} q & for \quad x=0 \\ p & for \quad x=1 \\ 0 & OW. \end{cases} \qquad ;\forall x\in\mathbb{R}$$

٣- بالعودة إلى المثال / ٣/ من (٣ ,١ ,١ ,٥) نجد أنَّ قانون توزيع المتغيِّر العشوائي $X$ يُعيَّن على النحو الآتي. من أجل أي حادث $B$ من $\Re$ لدينا:

$$P_X(B)=P\left(X^{-1}(B)\right)=\sum_{\omega_1,\omega_2=1}^{6}\frac{1}{36}\,\delta_{(\omega_1,\omega_2)}\left(X^{-1}(B)\right)$$

وهذا يعني أنَّ:

$$P_X=\sum_{\omega_1,\omega_2=1}^{6}\frac{1}{36}\,\delta_{(\omega_1,\omega_2)}\circ X^{-1}$$

أو أن نقدِّمه بالشكل الآتي أيضاً. من أجل كل $x\in\mathbb{R}$ لدينا:

$$P_X\left(\{x\}\right)=P\left(X^{-1}\left(\{x\}\right)\right)=\sum_{\omega_1,\omega_2=1}^{6}\frac{1}{36}\,\delta_{(\omega_1,\omega_2)}\left(X^{-1}\left(\{x\}\right)\right)=\frac{1}{36}\sum_{\omega_1,\omega_2=1}^{6}\mathbf{I}_{X^{-1}(\{x\})}(\omega_1,\omega_2)$$

٤- بالعودة إلى المثال / ٤/ من (٣ ,١ ,١ ,٥) نجد أنَّ قانون توزيع المتغيِّر العشوائي $X$ يُعيَّن على النحو الآتي:

$$P_X(A)=\frac{1}{b-a}\int_{A\cap[a,b]}dx \qquad ;\forall A\in\Re$$

٥- ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً معطى على النحو الآتي:

$$\Omega=[0\,,\infty)\times[0\,,\infty)$$

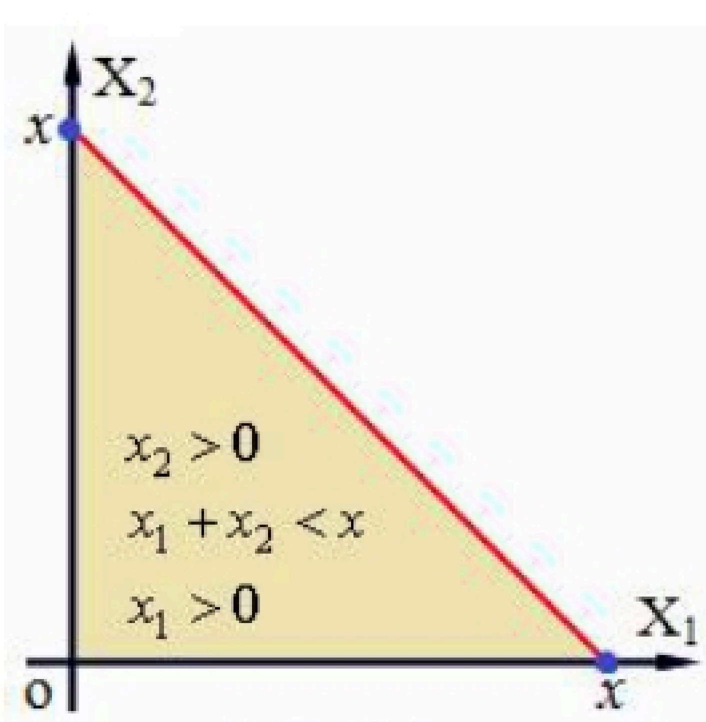
$$\mathscr{A}=\Re^2\cap[0\,,\infty)\times[0\,,\infty)$$

$$P(A):=\int_A \mathbf{e}^{-x_1}\,\mathbf{e}^{-x_2}\,dx_1\,dx_2 \qquad ;\forall A\in\Re$$

وليكن $X$ متغيِّراً عشوائياً فوق $[\Omega,\mathscr{A},P]$ مُعرَّفاً كما يلي:

$$X\left((x_1,x_2)\right):=x_1+x_2$$

والشكل الجانبي (٥ ,٢) يوضِّح مجموعة قيم المتغيِّر العشوائي $X$.



الشكل (٥ ,٢)

والآن لتعيين قانون توزيع $X$ سنأخذ $B\in\Re$ كيفية، فيكون لدينا الآتي:

$$P_X(B)=P\left(X^{-1}(B)\right)=\int_{X^{-1}(B)}P(d\omega)=\int P(d\omega)\,\mathbf{I}_B\left(X(\omega)\right)$$

$$=\int \mathbf{e}^{-(x_1+x_2)}\,\mathbf{I}_B\left(x_1+x_2\right)\,\lambda\left(d(x_1,x_2)\right)=\int \mathbf{e}^{-(x_1+x_2)}\,\mathbf{I}_B\left(x_1+x_2\right)dx_1\,dx_2$$

علماً أنَّ $\lambda$ هو قياس لوبيغ على $\Re^2$، وبما أنَّ الدالَّة المستكملة قابلة للمكاملة بحسب مفهوم ريمان، فإنَّه يمكننا أن نعيِّن $P_X$ من خلال العلاقة الآتية:

$$P_X(B)=\int \mathbf{e}^{-(x_1+x_2)}\,\mathbf{I}_B\left(x_1+x_2\right)dx_1\,dx_2$$

### (٥,٢,٢) دالَّة التوزيع لمتغيِّر عشوائي Distribution Function of Random Variable

في كثير من الحالات، وخاصّة لدى الجوانب التطبيقية والعمليّة، لا يظهر فيها المتغيِّر العشوائي بشكله الصريح؛ ولذا يستعاض عنه بتقديم ما يُعْرف باسم دالَّة توزيع Distribution Function متغيِّر عشوائي (ويطلق عليها البعض اسم دالَّة التوزيع التراكمية Cumulative Distribution Function أيضا)، وهذه الدالَّة تساعدنا في حساب احتمالات متعلِّقة بمتغيِّر عشوائي. إنَّ التعريف الرياضياتي لهذا المفهوم تقدّمه لنا الفقرة الآتية.

#### (٥,٢,٢,١) تعريف (دالَّة توزيع متغيِّر عشوائي)

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، ولنعرّف من أجل هذا المتغيِّر العشوائي دالَّة حقيقية $F_X$ على $\mathbb{R}$ من خلال العلاقة الآتية:

$$F_X(x)=P\left(\{\omega\in\Omega\,;X(\omega)<x\}\right)\qquad;\forall x\in\mathbb{R} \tag{5,8}$$

عندئذ تُدعى $F_X$ دالَّة التوزيع للمتغيِّر العشوائي $X$.

#### (٥,٢,٢,٢) ملاحظات

١- من أجل أي $x$ من $\mathbb{R}$ نجد أنَّ:

$$F_X(x)=P\left(\{\omega\in\Omega\,;X(\omega)<x\}\right)=P\left(X^{-1}(-\infty,x)\right)=P_X\left((-\infty,x)\right)$$

وهذا يعني أنَّ قيمة دالَّة التوزيع الاحتمالية $F_X$ عند نقطة $x$ من $\mathbb{R}$ ما هي إلاَّ احتمال الحادث $(-\infty,x)$ وفقاً لقانون توزيع المتغيِّر العشوائي $X$.

٢- كما هو واضح من تعريف دالَّة التوزيع الاحتمالية $F_X$ فإنَّ مجموعة قيمها هي الفترة $[0,1]$.

٣- إنَّ تعريف دالَّة توزيع متغيِّر عشوائي ما هو إلا حالة خاصة من تعريف أكثر شمولية يُعرف باسم دالَّة توزيع قياس احتمالي الذي يُصاغ على النحو الآتي: لنأخذ $Q$ قياساً احتمالياً على الفضاء المقيس $[\mathbb{R},\mathfrak{R}]$، ولنعرِّف من أجله دالَّةً حقيقيةً $F_Q$ على $\mathbb{R}$ من خلال العلاقة الآتية:

$$F_Q(x):=Q\left((-\infty,x)\right)\qquad;\forall x\in\mathbb{R} \tag{5,9}$$

فعندئذ الدالَّة $F_Q$ تُدعى دالَّة توزيع القياس الاحتمالي $Q$، وهنا يجب الانتباه إلى أنَّه في الحالة العامّة لدينا $Q\left([a,b)\right)\neq Q\left([a,b]\right)$ وذلك من أجل $-\infty<a<b<+\infty$.

الآن لو افترضنا أنَّ $X$ متغيِّر عشوائي فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، وقمنا بأخذ $Q=P_X$ فإنَّه بحسب الملاحظة السّابقة سنحصل على دالَّة توزيع $X$ أيضاً، وذلك لأنَّه من أجل أيّة قيمة $x\ni\mathbb{R}$ لدينا ما يلي:

$$F_{P_X}(x)=P_X\left((-\infty\;,\;x)\right)=P\left(X^{-1}(-\infty\;,\;x)\right)=P\left(X\in(-\infty\;,\;x)\right)=P\left(X<x\right)=F_X(x)$$

فيما يلي سنقدّم بعض المبرهنات التي تُظهر لنا بعض خصائص دالَّة التوزيع الاحتمالية لمتغيِّر عشوائي واستخداماتها، وسنبدأها بالمبرهنة الآتية التي تُمكننا من استخدام دالَّة التوزيع الاحتمالية في حساب احتمالات تتعلّق بمتغيّر عشوائي $X$.

#### (٥,٢,٢,٣) مبرهنة

لتكن $F_X$ دالَّة توزيع متغيِّر عشوائي $X$ معرَّف فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فعندئذ من أجل $-\infty<a<b<+\infty$ يكون لدينا العلاقة الآتية محقَّقةً:

$$P(a\le X<b)=F_X(b)-F_X(a) \tag{5,10}$$

**البرهان**: بما أنَّ $a<b$ فعندئذ سيكون لدينا:

$$\{\omega \in \Omega \;;\; X(\omega) < b\} \supset \{\omega \in \Omega \;;\; X(\omega) < a\}$$

ومن ثمَّ يمكننا أن نكتب الآتي:

$$\begin{aligned} P(a \le X < b) &= P\left(\{\omega\in\Omega\,;\, a \le X(\omega) < b\}\right) = P\left(\{\omega\in\Omega\,;\, X(\omega)<b\} \setminus \{\omega\in\Omega\,;\, X(\omega)<a\}\right) \\ &= P\left(\{\omega\in\Omega\,;\, X(\omega)<b\}\right) - P\left(\{\omega\in\Omega\,;\, X(\omega)<a\}\right) \\ &= P(X<b) - P(X<a) = F_X(b) - F_X(a) \end{aligned}$$

### (٥,٢,٣) الخصائص المميِّزة لدوال التوزيع الاحتمالية على $\mathbb{R}$

المبرهنة الآتية تقدِّم لنا أهمّ الخصائص المميزة لدالَّة توزيع احتمالية على $\mathbb{R}$ .

#### (٥,٢,٣,١) مبرهنة

لتكن $F_X$ دالَّة توزيع متغيِّر عشوائي $X$ معرَّف فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فعندئـذ يكون للدالَّة $F_X$ الخصائص الآتية:

$\mathbf{e_1}$ – الدالَّة $F_X$ غير متناقصة على $\mathbb{R}$ .

$\mathbf{e_2}$ – لدينا:

$$s_1)\quad \lim_{x\to-\infty} F_X(x) = 0 \qquad [5,11\text{-a}]$$

$$s_2)\quad \lim_{x\to+\infty} F_X(x) = 1 \qquad [5,11\text{-b}]$$

$\mathbf{e_3}$ – الدالَّة $F_X$ مستمرَّة من اليسار (على الأقل) عند كل نقطة $x \ni \mathbb{R}$ .

إنَّ هذه الخاصّية الأخيرة تكون موضع اهتمامنا عندما تعاني دالَّة التوزيع $F_X$ من نقاط انقطاع فقط، وذلك لأنَّه إذا كانت $F_X$ مستمرَّة على $\mathbb{R}$ فإنَّها تكون مستمرَّة من اليسار واليمين عند كل نقطة من $\mathbb{R}$ .

**البرهان**: من أجل الفقرة:

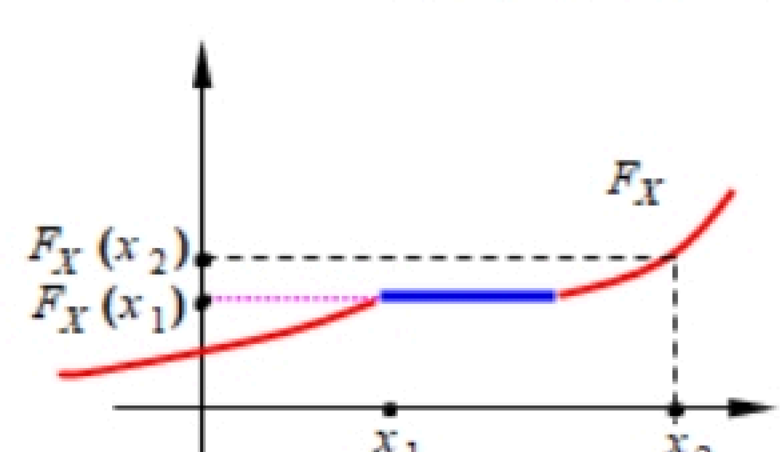
$\mathbf{e_1}$ – ليكن لدينا $x_1$ و $x_2$ من $\mathbb{R}$ مع $x_2 > x_1$، فعندئذ يكون لدينا:

$$\{\omega \in \Omega \;;\; X(\omega) < x_1\} \supset \{\omega \in \Omega \;;\; X(\omega) < x_2\}$$

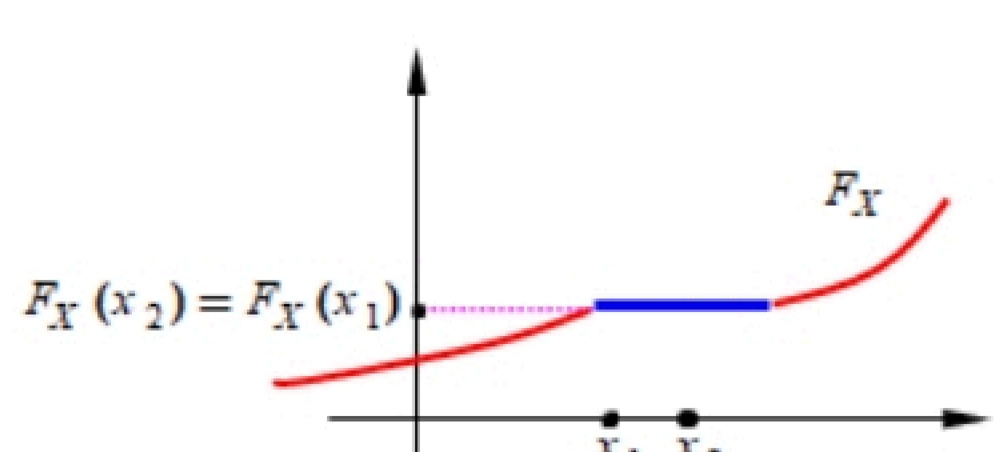
ومن ثمَّ بحسب خاصّية الاطراد للقياس الاحتمالي يمكننا أن نكتب الآتي:

$$F_X(x_2) = P\left(\{\omega \in \Omega \;;\; X(\omega) < x_2\}\right) \ge P\left(\{\omega \in \Omega \;;\; X(\omega) < x_1\}\right) = F_X(x_1)$$

وهذا يعني أنَّ الدالَّة $F_X$ غير متناقصة على $\mathbb{R}$ . إنَّ حالة المساواة $F_X(x_2) = F_X(x_1)$ تتحقَّق عندما تكون الدالَّة $F_X$ ثابتة على فترة تحوي $x_1$ و $x_2$ (انظر الشكلين الآتيين).




الشكل (٥,٣ أ)




الشكل (٥,٣ ب)

$\mathbf{e}_2$ – لإثبات صحة النهاية $(s_1)$ سنأخذ متتالية عددية $(x_n)_{n\in\mathbb{N}}$ متناقصة نحو $-\infty$ (أي إنَّ $\lim_{n\to\infty} x_n = -\infty$)، فعندئذ متتالية الحوادث $(-\infty, x_n)$ سوف تتقارب من الأعلى من الحادث المستحيل $\varnothing$، أي إنَّ $(-\infty, x_n) \underset{n\to\infty}{\downarrow} \varnothing$، ومن ثمَّ بحسب خاصية الاستمرار من الأعلى للقياسات يمكننا أن نكتب الآتي:

$$\lim_{x\to-\infty} F_X(x) = \lim_{n\to\infty} P_X\left((-\infty,x_n)\right) = P_X\left(\lim_{n\to\infty}(-\infty,x_n)\right) = P_X(\varnothing) = 0$$

وأمّا لإثبات صحة النهاية $(s_2)$ سنأخذ متتالية عددية $(y_n)_{n\in\mathbb{N}}$ متزايدة نحو $+\infty$ (أي إنَّ $\lim_{n\to\infty} y_n = \infty$)، فعندئذ متتالية الحوادث $(-\infty, y_n)$ سوف تتقارب من الأدنى نحو الحادث الأكيد $\mathbb{R}$، أي إنَّ $(-\infty, x_n) \underset{n\to\infty}{\uparrow} \mathbb{R}$، ومن ثمَّ بحسب خاصية الاستمرار من الأدنى للقياسات يمكننا أن نكتب الآتي:

$$\lim_{x\to+\infty} F_X(x) = \lim_{n\to\infty} P_X\left((-\infty,x_n)\right) = P_X\left(\lim_{n\to\infty}(-\infty,x_n)\right) = P_X(\mathbb{R}) = 1$$

$\mathbf{e}_3$ – لإثبات خاصية الاستمرار من اليسار للدالَّة $F_X$ عند نقاط الانقطاع سنأخذ $x_o \in \mathbb{R}$ نقطةً ما من نقط انقطاع الدالَّة $F_X$، ولنأخذ متتالية عددية $(y_n)_{n\in\mathbb{N}}$ متزايدة ومتقاربة من $x_o$، أي إنَّ $\lim_{n\to\infty} y_n = x_o$، فعندئذ متتالية الحوادث $(-\infty, y_n)$ سوف تتقارب من الأدنى نحو الحادث $(-\infty, x_o)$، ولذلك سيكون لدينا $\bigcup_{n=1}^{\infty}(-\infty, y_n) = (-\infty, x_o)$، ومن ثمَّ بحسب خاصية الاستمرار من الأدنى للقياسات ينتج لدينا:

$$\begin{aligned}
\lim_{n\to\infty} F_X(y_n) &= \lim_{n\to\infty} P_X\left((-\infty, y_n)\right) = P_X\left(\lim_{n\to\infty}(-\infty, y_n)\right) \\
&= P_X\left(\bigcup_{n=1}^{\infty}(-\infty, y_n)\right) = P_X\left((-\infty, x_o)\right) \\
&= P\left(X^{-1}(-\infty, x_o)\right) = P(X < x_o) = F_X(x_o)
\end{aligned}$$

وبهذا يتمّ البرهان.

### (٢, ٣, ٢, ٥) ملاحظة

يمكن البرهان (انظر [64]) على أنّه لكل دالَّة حقيقية $F$ معرَّفة على $\mathbb{R}$ وتتمتَّع بالخواص $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ يوجد قياس احتمالي وحيد $Q$ على $[\mathbb{R}, \Re]$ بحيث يكون لدينا ما يلي محُقَّقاً:

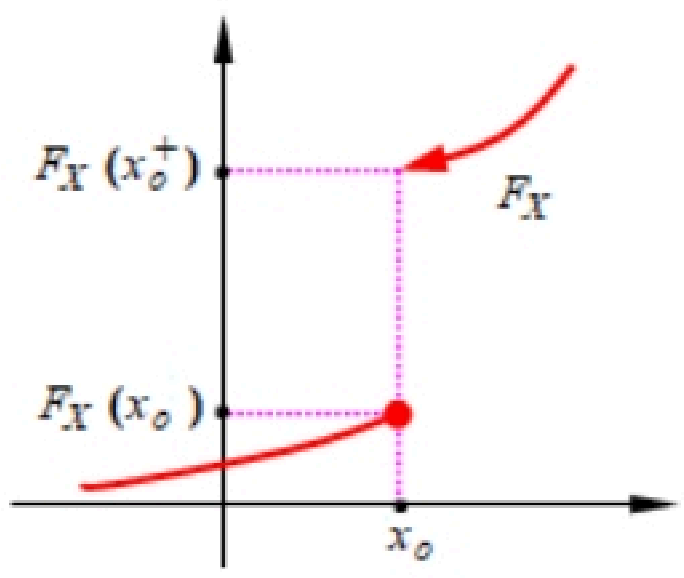
$$F_Q(x) = F(x) \qquad ; \forall x \in \mathbb{R} \qquad [5,12]$$

وهذا يعني أنَّ كل دالَّة حقيقية $F$ مُعرَّفة على $\mathbb{R}$ وتُحقِّق الخواص $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ هي دالَّة توزيع احتمالية، بمعنى أنَّه يوجد متغيِّر عشوائي $X$ بحيث يكون من أجل كل $x \in \mathbb{R}$ لدينا $F(x) = F_X(x)$ (من أجل ذلك يكفي أن نأخذ $Q = P_X$).

### (٢, ٣, ٣, ٥) نتائج

١- ممّا سبق ينتج لدينا أنَّ الخواص $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ تُميِّز دالَّة التوزيع الاحتمالية على $\mathbb{R}$ بشكل تام.

٢- إذا عانت دالَّة التوزيع الاحتمالية $F_X$ من نقطة انقطاع في موضع $x_o \in \mathbb{R}$ فإنَّها تقوم بقفزة نحو الأعلى عند هذه النقطة (انظر الشكل (٤, ٥))، وبما أنَّ دوال التوزيع الاحتمالية مستمرَّة من اليسار عند النقطة $x_o$ فإنَّه يوضع نقطة مضخَّمة على الطرف الأيمن لمنحنى الدالَّة ليدلِّل على قيمتها عند هذه النقطة في حين يوضع رأس سهم على الطرف الأيسر لمنحنى الدالَّة عند هذه النقطة للدلالة على عدم مساواة قيمة الدالَّة لهذا الموضع، والشكل الآتي يوضِّح ذلك، علماً أنَّ $F_X(x_o^+)$ تشير إلى قيمة النهاية من اليمين عند النقطة $x_o$.



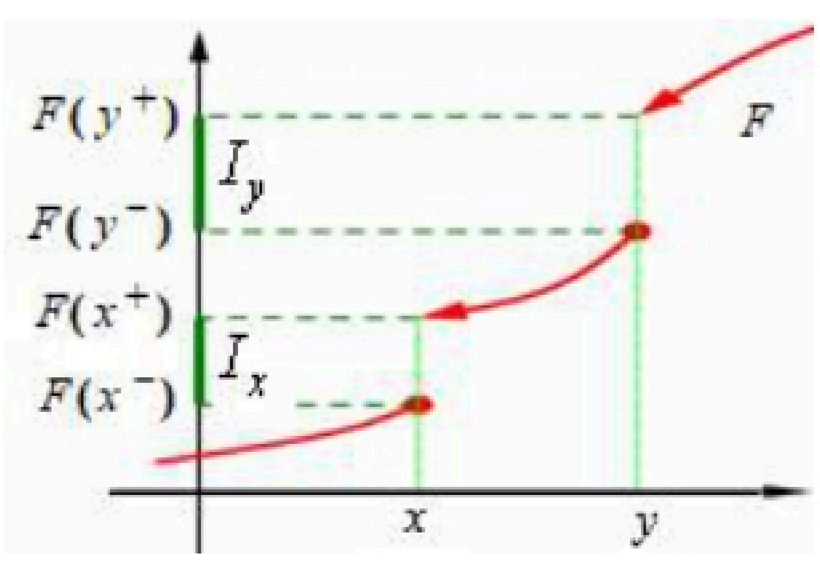

الشكل (٤ , ٥)

٣- إنَّ مجموعة نقط الانقطاع لدالَّة توزيع احتمالية $F$ تكون إمَّا منتهية أو قابلة للعدِّ على الأكثر.

**الإثبات:** بفرض أنَّ $x$ نقطة انقطاع لـ $F$ فإنَّه يمكننا من أجلها تعريف فترة مفتوحة $I_x$ كما يلي:

$$I_x = \left(F(x^-)\,,\,F(x^+)\right)$$

علماً أنَّ $F_X(x^-)$ و $F_X(x^+)$ تشيران على الترتيب إلى قيمة النهاية من اليسار واليمين للدالَّة $F$ عند النقطة $x$، فإذا كانت $x$ و $y$ نقطتي انقطاع مختلفتين لدالَّة التوزيع $F$، فعندئذ سيكون لدينا $I_x \cap I_y = \varnothing$ (انظر الشكل الآتي (٥ , ٥)) علماً أنَّ الفترة $I_y$ هي الفترة الناتجة عن نقطتي النهاية $F(y^-)$ و $F(y^+)$، أي إنَّ $I_y = \left(F(y^-)\,,\,F(y^+)\right)$.




الشكل (٥ , ٥)

و من جهة أخرى نعلم (من نظرية المجموعات) أنَّ مجموعة كل الفترات المفتوحة في $\mathbb{R}$ تكون منتهية أو قابلة للعدِّ على الأكثر، ومن ثمَّ الفترة $[0,1]$ (التي تمثِّل مجموعة قيم الدالَّة $F$) ستحتوي على الأكثر عدداً قابلاً للعدِّ من الفترات المفتوحة النّاتجة عن نقط الانقطاع لـ $F$، وذلك لأنَّ الفترة $[0,1]$ تحوي كل الفترات المفتوحة النّاتجة عن نقط الانقطاع لـ $F$.

### (٤ , ٣ , ٢ , ٥) ملاحظات

١- إنَّ مجموعة نقط الاستمرار لدالَّة توزيع احتمالية $F$ هي مجموعة كثيفة في $\mathbb{R}$ دوماً، ومع ذلك فمن الممكن أنَّ تكون مجموعة نقط الانقطاع لدالَّة التوزيع الاحتمالية $F$ هي مجموعة كثيفة في $\mathbb{R}$ أيضاً، ولتوضيح ذلك سنأخذ $X$ متغيِّرا عشوائيًّا مجموعة قيمه $\mathbf{X} = \{x_n \in \mathbb{Q}\ ;\ \forall n \in \mathbb{N}\}$، وقانون توزيعه مُعطى من خلال العلاقة الآتية:

$$P_X(\{x_n\}) = P(X = x_n) = \frac{6}{\pi^2 n^2} \qquad ; n \in \mathbb{N}$$

فعندئذ سيكون لدينا:

$$F_X(x) = \sum_{\substack{n=1 \\ x_n < x}}^{\infty} P(X = x_n)$$

فنلاحظ أنَّ مجموعة نقط الانقطاع للدالَّة $F_X$ هي مجموعة الأعداد العادية $\mathbb{Q}$، وهي مجموعة كثيفة في $\mathbb{R}$ كما نعلم.

٢- تُعرِّف بعض المراجع دالَّة التوزيع لمتغيِّر عشوائي $X$ من خلال العلاقة الآتية:

$$F_X(x) = P(X \le x) \qquad ; \forall x \in \mathbb{R} \qquad [5,13]$$

وفي هذه الحالة تصبح دالَّة التوزيع $F_X$ مستمرَّة من اليمين (عوضاً عن الاستمرار من اليسار) عند كل نقطة $x \in \mathbb{R}$، وأكثر من ذلك ستزداد قيمة هذه الدالَّة في كل نقطة انقطاع $x_o$ بمقدار يساوي:

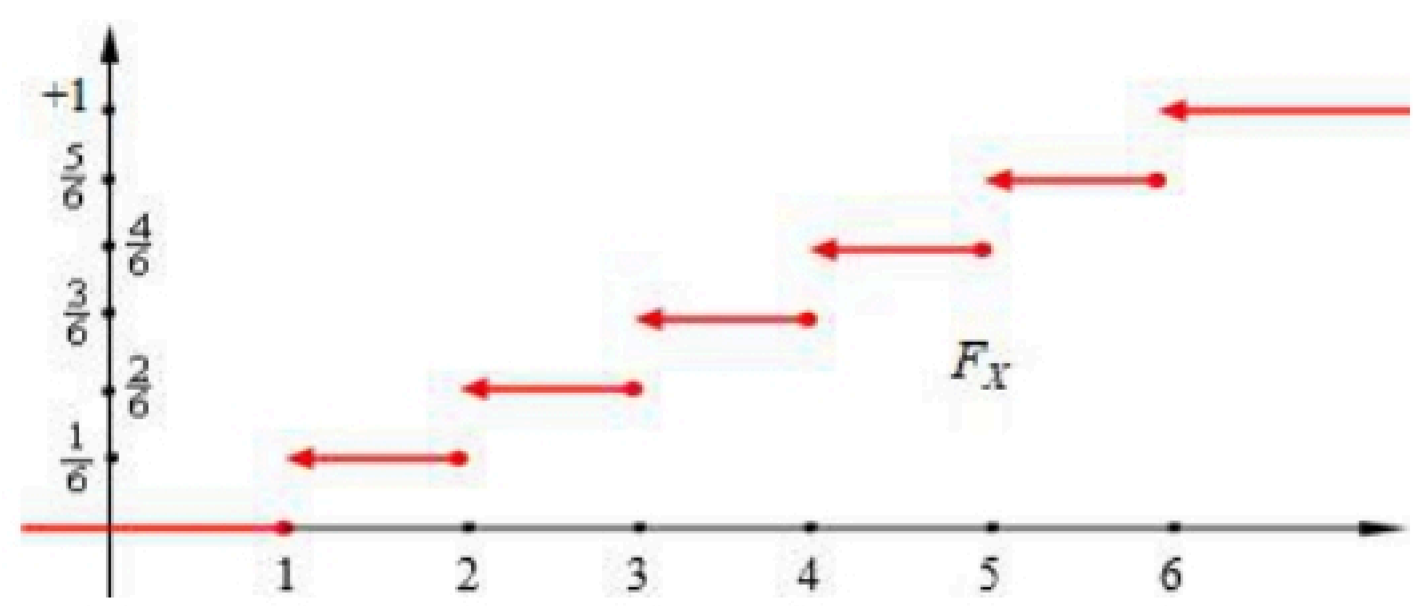
$$\Delta = F_X(x_o^+) - F_X(x_o^-) = P(X \le x_o) - P(X < x_o)$$

ولتوضيح ذلك لنأخذ تجربة إلقاء حجر نرد متوازن لمرّة واحدة فقط، ولنأخذ $X$ تطبيقاً فوق فوق الفضاء الاحتمالي لهذه التجربة العشوائيّة معرفاً من خلال العلاقة $X(\omega) = \omega$، فعندئذ نجد أنَّ $X$ هو متغيِّر عشوائي (تحقّق من صحة ذلك) متقطِّع؛ وذلك لأنَّ مجموعة قيمه

$\mathbf{X}=\{1,2,3,4,5,6\}$ منتهية، فإذا كانت دالَّة توزيع هذا المتغيِّر العشوائي معرَّفةً بالعلاقة $F_X(x)=P(X<x)$ ، فإنَّه سيكون لدالَّة توزيع هذا المتغيِّر العشوائي العرض الآتي:

$$F_X(x)=\begin{cases}0 & for\ x\le 1\\ \frac{i}{6} & for\ i<x\le i+1 \quad for\ \ i=1,2,3,4,5\\ 1 & for\quad x>6\end{cases}$$

ولرسمها البياني الشكل الآتي:




الشكل (٦, ٥ .أ)

فيكون لدينا على سبيل المثال ما يلي:

$$F_X(1)=0 \quad \& \quad F_X(1+\alpha)=\frac{1}{6} \qquad ;\ 0<\alpha\le 1$$

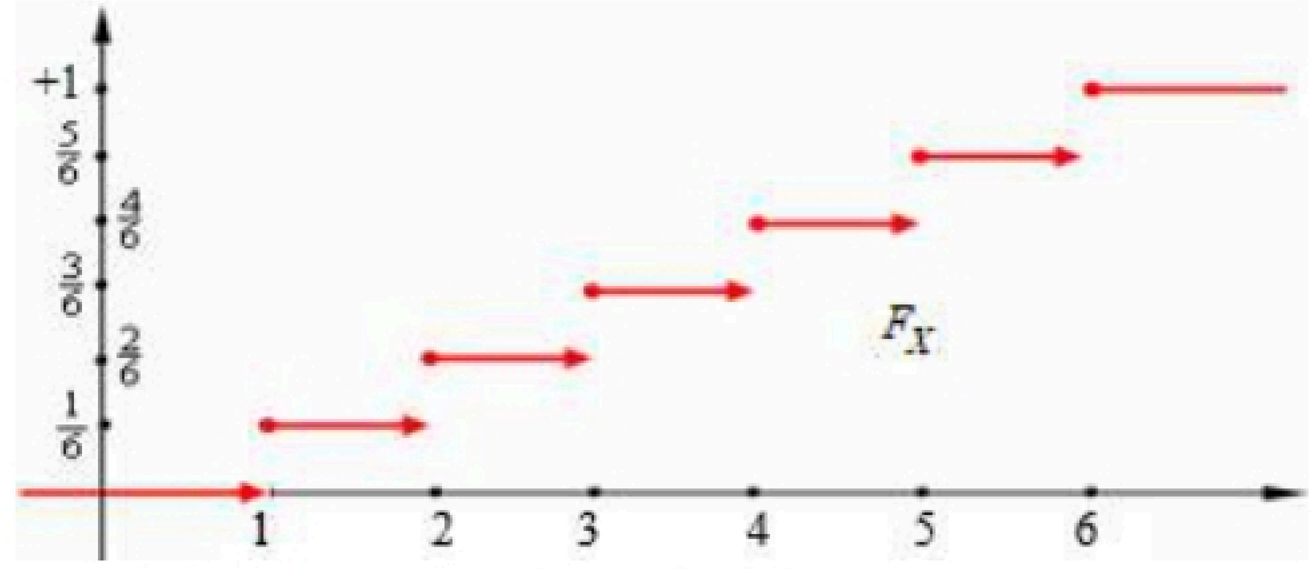
$$F_X(3)=\frac{2}{6} \quad \& \quad F_X(3+\alpha)=\frac{3}{6} \qquad ;\ 0<\alpha\le 1$$

$$F_X(6)=\frac{5}{6} \quad \& \quad F_X(6+\alpha)=1 \qquad ;\ \alpha>0$$

وأمَّا إذا كانت دالَّة توزيع المتغيِّر العشوائي $X$ معرَّفةً بالعلاقة [5,13] ، فعندئذ سيكون لدالَّة توزيع هذا المتغيِّر العشوائي العرض الآتي:

$$F_X(x)=\begin{cases}0 & for\ x<1\\ \frac{i}{6} & for\ i\le x<i+1 \quad ;\ i=1,2,\dots,5\\ 1 & for\quad x\ge 6\end{cases}$$

و أمَّا رسمها البياني فله الشكل الآتي:




الشكل (٦, ٥ .ب)

ويكون لدينا في هذه الحالة (على سبيل المثال) ما يلي:

$$F_X(1)=\frac{1}{6} \quad \& \quad F_X(1-\alpha)=0 \qquad ;\alpha>0$$

$$F_X(3)=\frac{3}{6} \quad \& \quad F_X(3-\alpha)=\frac{2}{6} \qquad ;\ 0<\alpha\le 1$$

$$F_X(6)=1 \quad \& \quad F_X(6-\alpha)=\frac{5}{6} \qquad ;\ 0<\alpha\le 1$$

**(٤ ,٢ ,٥) مبرهنة (الشرط اللازم والكافي لاستمرار دالَّة توزيع احتمالية على $\mathbb{R}$ )**

لتكن $F_X$ دالَّة توزيع متغيِّر عشوائي $X$ فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$، فعندئذ تكون الدالَّة $F_X$ مستمرَّة على $\mathbb{R}$ **إذا وفقط إذا** كان من أجل كل $x\in\mathbb{R}$ لدينا العلاقة الآتية مُحَقَّقةً:

$$P(X=x)=0 \qquad [5,14]$$

**البرهان: مــن أجل لزوم الشرط:**

لدينا الدالَّة $F_X$ مستمرَّة على $\mathbb{R}$، ومن ثَمَّ بحسب تعريف الاستمرار للدوال يكون من أجل كل $\varepsilon>0$ يوجد $\delta>0$ بحيث أنَّه من أجل $x\in\mathbb{R}$ نقطة استمرار لـ $F_X$، ومن أجل كل $y\in\mathbb{R}$ يكون لدينا $|F_X(y)-F_X(x)|<\varepsilon$ مادام $|y-x|<\delta$، فلو أخذنا $y>x$ (على سبيل المثال لا الحصر) فإنَّه بحسب خاصة عدم التناقص لدوال التوزيع الاحتمالية سيكون من أجل $|y-x|<\delta$ الآتي محققَّاً:

$$|F_X(y)-F_X(x)|=F_X(y)-F_X(x)=P(x\le X<y)$$

ومن جانب آخر نعلم من خاصية الاطراد للقياس الاحتمالي $P$ أنَّ:

$$P(X=x)\le P(x\le X<y)$$

ومن ثَمَّ ينتج لدينا $P_X(\{x\})<\varepsilon$ وذلك لأنَّ:

$$P(X=x)=P_X(\{x\})\le P(x\le X<y)<\varepsilon$$

وذلك بغض النظر عن قيمة $y\in\mathbb{R}$، ومنه بحكم إمكانية اختيار العدد $\varepsilon$ صغيراً بالقدر الذي نريد يمكننا أن نكتب $P(X=x)=0$ وبهذا يتمَ الإثبات.

**مــن أجل كفاية الشرط:**

لدينا $P(X=x)=0$ من أجل كل $x\in\mathbb{R}$، فعندئذ لإثبات استمرار الدالَّة $F_X$ يكفي أن نثبت أنـَّها مستمرَّة من اليمين عند كل نقطة $x\in\mathbb{R}$ (لأنها مستمرَّة من اليسار عند كل نقطة $x\in\mathbb{R}$). لذلك لنأخذ $(y_n)_{n\in\mathbb{N}}$ متتالية عددية بحيث تكون $x<y_n$ من أجل كل $n\in\mathbb{N}$ وكذلك $\lim\limits_{n\to\infty} y_n=x$، فعندئذ بحسب مبرهنة في التحليل الرياضياتي حول الفترات المتداخلة، ومن خصائص دالَّة التوزيع $F_X$ والقياس الاحتمالي $P$ يُمكننا أن نكتب الآتي:

$$\begin{aligned}\lim_{n\to\infty} F_X(y_n)&=\lim_{n\to\infty} P_X\left((-\infty,y_n)\right)=P_X\left(\lim_{n\to\infty}(-\infty,y_n)\right)\\&=P_X\left((-\infty,x]\right)=P_X\left((-\infty,x)\right)+P_X\left(\{x\}\right)=F_X(x)+P_X\left(\{x\}\right)\end{aligned}$$

ولكن من الفرض لدينا $P_X(\{x\})=0$ من أجل كل $x\in\mathbb{R}$، ولذلك يكون لدينا:

$$\lim_{n\to\infty} F_X(y_n)=F_X(x) \qquad ;\ \forall x\in\mathbb{R}$$

وهذا يعني أنَّ دالَّة التوزيع $F_X$ مستمرَّة من اليمين عند كل نقطة $x\in\mathbb{R}$، ومن ثَمَّ الدالَّة $F_X$ مستمرَّة عند كل نقطة $x\in\mathbb{R}$، وبهذا يتمّ البرهان.

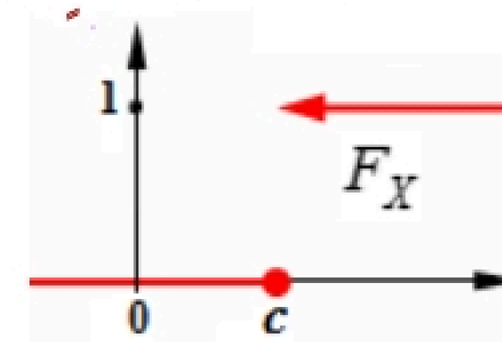
### (٥ , ٢ , ٤ , ١) أمثلة

١- بالعودة إلى المثال (٣ , ١ , ١ , ٥) حيث لدينا $X$ متغيِّر عشـوائي معرّف على $\Omega$ من خلال العلاقة الآتية:

$$X(\omega) = c \qquad ; \forall\, \omega \in \Omega$$

علماً أنَّ $c$ ثابت عددي، فعندئذ نجد أنَّ لدالَّة توزيع المتغيِّر العشوائي $X$ العرض الآتي:

$$F_X(x) = P(X < x) = \begin{cases} 0 & for \quad x \le c \\ 1 & for \quad x > c \end{cases}$$




الشكل (٥ , ٧) الرسم البياني لدالَّة التوزيع الاحتمالية.

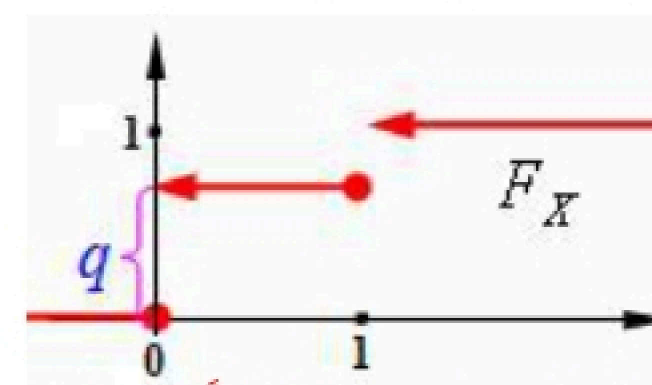
ونذكّر هنا مرَّةً أخرى أنَّ الدائرة المغلقة عند نقطة الانقطاع تدلّ على النهاية من اليسار عند هذه النقطة والمساوية لـ $F_X(c)$ ، بينما يدل رأس السهم عند نقطة الانقطاع إلى قيمة النهاية من اليمين عند هذه النقطة. كذلك نلاحظ أنَّ لهذه الدالَّة شكل دالَّة درجية بقفزة واحدة فقط عند $c$ ، ولذا فشكلها البياني هو أبسط شكل بين الأشكال البيانية للدوال الدرجية.

٢- بالعودة إلى المثال / ٢/ من (٣ , ١ , ١ , ٥) حيث لدينا $X$ متغيِّراً عشوائياً معرّفاً على $\Omega$ من خلال العلاقة الآتية:

$$X(\omega) = \begin{cases} 0 & for \quad \omega = \omega_2 \\ 1 & for \quad \omega = \omega_1 \end{cases}$$

فلو افترضنا أنَّ $p := P(X = 1)$ ، فعندئذ نجد أنَّ دالَّة التوزيع الاحتمالية لهذا المتغيِّر العشوائي لها العرض الآتي:

$$F_X(x) = \begin{cases} 0 & for\ x \le 0 \\ 1 - p = q & for\ 0 < x \le 1 \\ 1 & for\ x > 1 \end{cases}$$



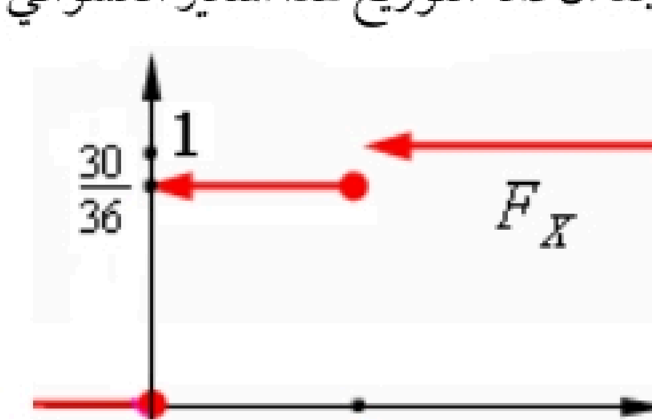

الشكل (٥ , ٨) الرسم البياني لدالَّة التوزيع الاحتمالية (لهذه الدالَّة شكل دالَّة درجية بقفزتين فقط).

٣- بالعودة إلى المثال / ٣/ من (٣ , ١ , ١ , ٥) حيث لدينا $X$ متغيِّر عشـوائي معرّف على $\Omega$ من خلال العلاقة الآتية:

$$X(\omega) = \begin{cases} 0 & for \quad \omega_1 + \omega_2 < 10 \\ 1 & for \quad \omega_1 + \omega_2 \ge 10 \end{cases}$$

فنجد أنَّ دالَّة التوزيع لهذا المتغيِّر العشوائي لها العرض والشكل الآتيين:

$$F_X(x) = P(X < x) = \begin{cases} 0 & for\ x \le 0 \\ \frac{30}{36} & for\ 0 < x \le 1 \\ 1 & for\ x > 1 \end{cases}$$



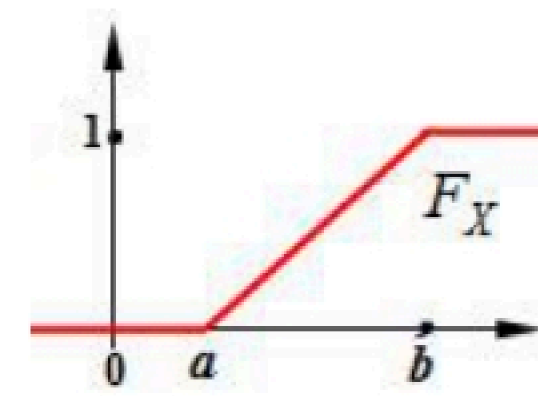

الشكل (٥ , ٩) الرسم البياني لدالَّة التوزيع الاحتمالية.

٤– بالعودة إلى المثال / ٤/ من (٣ , ١ , ١ , ٥) حيث المتغيِّر العشوائي $X$ مُعرَّف من خلال العلاقة $X(\omega)=\omega$ ، والفضاء الاحتمالي $[\Omega,\mathscr{A},P]$ مُعطى كما يلي: $\mathscr{A}$

$$\Omega=[a\,,b] \qquad \& \qquad \mathscr{A}=\Re\bigcap[a\,,b] \qquad \& \qquad P(A)=\frac{1}{b-a}\int_A dx$$

فنجد بحسابات بسيطة أنَّ لدالَّة توزيع هذا المتغيِّر العشوائي العرض الآتي:

$$F_X(x)=\begin{cases}0 & for \quad x\le a\\ \dfrac{x-a}{b-a} & for \quad a<x\le b\\ 1 & for \quad x>b\end{cases}$$




الشكل (٥ , ١٠) الرسم البياني لدالَّة التوزيع الاحتمالية (دالَّة مستمرة على $\mathbb{R}$).

وهنا نلاحظ أنَّ دالَّة التوزيع الاحتمالية $F_X$ مستمرَّة على $\mathbb{R}$ ، في حين أنَّ دوال التوزيع الاحتمالية في الأمثلة الثلاث السابقة كانت دوال درجية (هذه الملحوظة توحي لنا إلى وجود تصنيفات للمتغيِّرات العشوائية).

٥– بالعودة إلى المثال / ٥/ من (٣ , ١ , ١ , ٥) حيث المتغيِّر العشوائي $X$ معرّف على $\Omega$ من خلال العلاقة الآتية:

$$X\left((x_1,x_2)\right):=x_1+x_2$$

فنجد من أجل كل $x>0$ أنَّ:

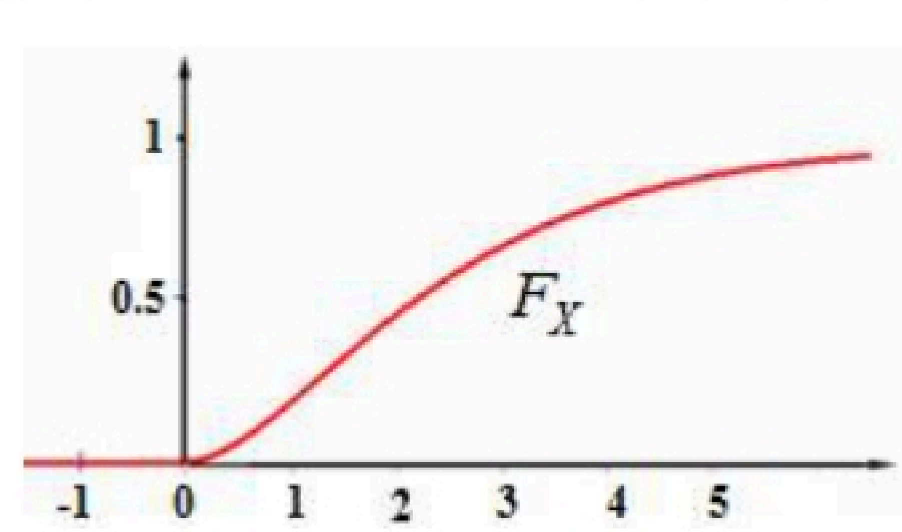
$$F_X(x)=P\left(\left\{(x_1,x_2)\in\Omega\;\middle|\;X\left((x_1\,,\,x_2)\right)<x\right\}\right)$$
$$=\iint\limits_{x_1,x_2\ge0;\,x_1+x_2<x}\mathbf{e}^{-x_1}\mathbf{e}^{-x_2}\,dx_1\,dx_2=\int\limits_{x_1=0}^{x}\;\int\limits_{x_2=0}^{x-x_1}\mathbf{e}^{-x_1}\mathbf{e}^{-x_2}\,dx_1\,dx_2$$

وبحساب التكامل الذي في الطرف الأيمن من العلاقة الأخيرة نجد:

$$F_X(x)=1-\frac{x+1}{\mathbf{e}^x}$$

وأمّا من أجل كل $x\le0$ فإنَّنا سنضع $F_X(x)=0$ (قدِّم تبريراً لهذا الافترض)، وبذلك يكون لدالَّة توزيع $X$ العرض الآتي:

$$F_X(x)=\begin{cases}0 & for\ x\le0\\ 1-\dfrac{x+1}{\mathbf{e}^x} & for\ x>0\end{cases}$$




الشكل (٥ , ١١) الرسم البياني لدالَّة التوزيع الاحتمالية (دالَّة مستمرة على $\mathbb{R}$).

٦- لنقم بإلقاء قطعة نقود (ليست بالضـرورة أنَّ تكون متوازنة) لمرَّة واحدة فقط، فإذا حصلنا على صورة فإنَّنا نلق قضيباً على المستوى الأفقي بشكل عشوائي، فيصنع زاوية $\alpha$ مع الشمال الجغرافي، وأمَّا إذا حصلنا على شعار فإنَّنا ننهي التجربة. عندئذ سيكون لمجموعة نتائج هذه التّجربة العشوائيّة العرض الآتي:

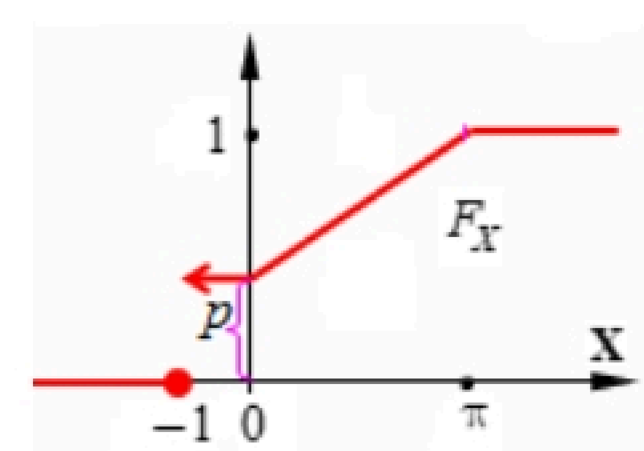
$$\Omega = \{T\} + \{(H , \alpha) ; \alpha \in [0 , \pi)\}$$

ولنعرّف متغيِّرا عشـوائيًا $X$ فوق الفضاء الاحتمالي لهذه التجربة من خلال العلاقة الآتية:

$$X(\omega) := \begin{cases} -1 & for \ \omega = T \\ \\ \alpha & for \ \omega = (H , \alpha) \end{cases}$$

فنجد أنَّ مجموعة قيم هذا المتغيِّر العشوائي هي $\mathbf{X} = \{-1\} \cup [0 , \pi)$ ، ومن ناحية أخرى بفرض أنَّ احتمال الحصول على الشعار يساوي $0 < p < 1$ ، فإنَّه سيكون لدالَّة التوزيع الاحتمالية لهذا المتغيِّر العشـوائي العرض الآتي:

$$F_X(x) = \begin{cases} 0 & for \ x \le -1 \\ p & for \ -1 < x \le 0 \\ p + \dfrac{1-p}{\pi} x & for \ 0 < x \le \pi \\ 1 & for \ x > \pi \end{cases}$$




الشكل(١٢ , ٥) الرسم البياني لدالَّة التوزيع الاحتمالية (دالَّة ليست مستمرّة على R).

فنلاحظ من الرسم البياني للدالَّة $F_X$ أنّها مستمرَّة على المجموعة $\mathbb{R} \setminus \{-1\}$ حيث تعاني من نقطة انقطاع عند النقطة $x = -1$ .

## (٥,٣) تصنيف المتغيرات العشوائية
## Classification of Random Variables

لو أمعنا النظر في الأمثلة السّابقة (١ , ٤ , ٢ , ٥) فإنَّنا نلاحظ أنَّ هناك متغيِّرات عشوائية دوال توزيعها درجية كما في الأمثلة / ١ و٢ و٣/ ، وأخرى دوال توزيعها مستمرَّة كما في المثالين / ٤ و٥/ ، ونوع ثالث دوال توزيعها ليست درجية وليست مستمرَّة كما في المثال / ٦/ ، وبناءً على ذلك يمكننا تصنيف المتغيِّرات العشوائيّة في ثلاثة أصناف رئيسة هي:

١- متغيِّرات عشوائية دوال توزيعها الاحتمالية مستمرَّة على $\mathbb{R}$ ، وهذا النوع يُعرف باسم متغيِّرات عشوائية مستمرَّة Continuous Random Variables . إنَّ مجموعة قيم هذا النوع من المتغيِّرات العشوائية تكون غير قابلة للعدِّ دوماً.

٢- متغيِّرات عشوائية دوال توزيعها ليست مستمرَّة على $\mathbb{R}$ ، وهنا نميِّز بين نوعين من هذه المتغيِّرات العشوائيّة أيضاً هما:

أ- متغيِّرات عشوائية دوال توزيعها ليست مستمرّة على كامل $\mathbb{R}$ (تعاني من نقطة انقطاع واحدة على الأقل)، ولكنَّها ليست دوال درجية أيضاً، فعندئذ يُقال عن هذا النوع من المتغيِّرات العشوائيّة إنّها متغيِّرات عشوائيّة غير مسـتمرة Uncontinuous Random Variables، ومن الأمثلة على هذا النوع المثال / ٦/ من الفقرة (١ , ٤ , ٢ , ٥) حيث عرض لنا متغيِّراً عشوائياً دالَّة توزيعه $F_X$ تعاني من انقطاع عند النقطة $x = -1$ ورسمها البياني بيّن أنَّها ليست دالَّة درجية. إنَّ مجموعة قيم هذا النوع من المتغيِّرات العشوائية ستكون مجموعات غير قابلة للعدِّ دوماً. (في كتابنا هذا لن نتطرق إلى دراسة هذا النوع من المتغيِّرات العشوائيّة).

ب- متغيِّرات عشوائية دوال توزيعها ليست مستمرّة على كامل $\mathbb{R}$، ولكنَّها دوال درجية كما في الأمثلة / ١ و٢ و٣/

من(١ , ٤ , ٢ , ٥) . إنَّ هذا النوع من المتغيرات العشوائيّة يُعرف باسم المتغيِّرات العشوائيّة المتقطِّعة Discrete Random Variables. إنَّ مجموعة القيم لهذا النوع من المتغيِّرات العشوائية تكون قابلة للعدِّ على الأكثر حتماً.

إذاً يمكن تصنيف المتغيِّرات العشوائيّة في ثلاثة أنواع رئيسة وفقاً لشكل دالَّة التوزيع، ولكن بما أنَّ دوال التوزيع تأتي كمرحلة لاحقة لتعيين المتغيِّر العشوائي فإنَّ ذلك يوجب علينا تقديم التصنيف انطلاقاً من الخصائص التي يتمتَّع بها المتغيِّر العشوائي نفسه بحيث تؤول دوال توزيعها إلى أحد الأصناف الثلاثة السّابقة.

نقدّم فيما يأتي تعاريف لهذه الأصناف الثلاثة، وسنركِّز (في دراستنا لهذا الكتاب) على المتغيِّرات العشوائيّة المتقطّعة والمستمرَّة فقط.

### (٥ , ٣ , ١) المتغيِّرات العشوائيّة المتقطِّعة Discrete Random Variables

نبدأ دراسة المتغيرات العشوائيّة المتقطّعة بالتعريف الآتي.

#### (٥ , ٣ , ١ , ١) تعريف (المتغيِّر العشوائي المتقطِّع)

ليكن $X$ متغيِّراً عشـوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فإذا كانت مجموعة قيم $X$ منتهيةً أو غير منتهية ولكنَّها قابلة للعدِّ، فعندئذ يُقال عن هذا المتغيِّر العشوائي إنَّه متقطِّع Discrete.

#### (٥ , ٣ , ١ , ٢) ملاحظات

١- يمكن عرض المتغيِّر العشوائي المتقطِّع على شكل دالَّة درجية، وذلك لأنَّه لو افترضنا $\mathbf{X}=\{x_i \;;\; i\in I\}$ هي مجموعة قيم متغيِّر عشـوائي متقطِّع $X$ حيث لدينا مجموعة الأدلة $I$ إمّا أن تكون منتهيةً أو غير منتهية ولكنَّها قابلة للعدّ، وقمنا بأخذ حوادث $Z_i$ لها العرض الآتي:

$$Z_i=\{\omega\in\Omega \;;\; X(\omega)=x_i\} \qquad ;\forall\, i=I$$

فإنَّنا نجد من أجل أيَّة قيمة $i \ni I$ أنَّ $Z_i \ni \mathscr{A}$، وكذلك $X^{-1}(\{x_i\})=Z_i$، وأكثر من ذلك نلاحظ أنَّ أسرة الحوادث $\mathscr{Z}=\{Z_i \;;\; i\in I\}$ تشكِّل تجزئةً للحادث الأكيد $\Omega$، ومن ثَمَّ يمكن كتابة المتغيِّر العشوائي المتقطِّع $X$ على النحو الآتي:

$$X(\omega)=\sum_{i\in I} x_i\cdot\mathbf{I}_{Z_i}(\omega) \qquad ;\forall\,\omega=\Omega \qquad [5,15\text{-a}]$$

وهذا يعني أنَّ لـ $X$ نموذج دالَّة درجية.

٢- إذا كانت مجموعة الأدلة $I$ منتهية، فعندئذ يمكننا أن نكتب $\mathbf{X}=\{x_i \;;\; i\in\mathbb{N}_n\}$ مع $n$ قيمة صحيحة مثَّبتة، ومن ثَمَّ يصبح لصيغة المتغيِّر العشوائي $X$ العرض الآتي:

$$X=\sum_{k=1}^{n} x_k\cdot\mathbf{I}_{Z_k} \qquad ;Z_1,Z_2,\ldots,Z_n\in\mathscr{A} \qquad [5,15\text{-b}]$$

وفي هذه الحالة يُقال عن المتغيّر العشوائي $X$ إنَّه بسيط Simple Random Variable، فعلى سبيل المثال لو أخذنا تجربة إلقاء حجر نرد متوازن لمرة واحدة، وبفرض أنَّ $X$ متغيِّر عشوائي فوق الفضاء الاحتمالي لهذه التجربة معرفاً من خلال العلاقة $X(\omega)=\omega$، فعندئذ يكون لهذا المتغيِّر العشوائي العرض الآتي:

$$X(\omega)=\sum_{k=1}^{6} k\cdot\mathbf{I}_{\{k\}}(\omega) \qquad ;\omega=1,2,3,4,5,6$$

ولذلك فإنَّ هذا المتغيِّر العشوائي بسيط.

٣- في حالة المتغيّرات العشوائيّة المتقطّعة يكون القياس $P_X$ معيَّناً وبشكل وحيد من خلال العلاقة الآتية:

$$P_X\left(\{x_i\}\right)=P\left(X=x_i\right) \qquad ;i\in I \qquad [5,16]$$

وذلك لأنَّه من أجل أيّة قيمة $i \in I$ لدينا:

$$P_X\left(\{x_i\}\right)=P\left(X^{-1}\left(\{x_i\}\right)\right)=P\left(Z_i\right)=P\left(\left\{\omega\in\Omega\ ;X(\omega)=x_i\ \right\}\right)=P\left(X=x_i\right)$$

ويُسمّى المقدار $P\left(X=x_i\right)$ **الكتلة الاحتمالية** Probability Mass للمتغيِّر العشوائي $X$، ويُرمز لها عادةً بـ $p_i$ من أجل أيّة قيمة $i\in I$ (كما قمنا باستخدامه سابقاً)، وفي بعض المراجع يُرمز لها بـ $f_X\left(x_i\right)$ ويدعونها **دالَّة الكتلة الاحتمالية** Probability Mass Function، أي إنَّه يُكتب:

$$p_i=P\left(X=x_i\right)=f_X\left(x_i\right) \qquad ;i\in I$$

لاحظ هنا أنَّه يجب أن يكون $\sum_{i\in I} p_i=1$.

**(٥,٣,٢,٣) أمثلة**

١- بالرجوع إلى الأمثلة /١ و ٢ و٣/ من الفقرة (٥,١,١,٣) نجد أنَّها متغيِّرات عشوائية متقطِّعة بمجموعة قيم منتهية، وكذلك الأمر بالنسبة إلى الدالَّة المميِّزة $\mathbf{I}_A$ عندما يكون $A$ حادثاً.

٢- بالرجوع إلى المثال /١/ المقدَّم في الملاحظة (٥,٢,٣,٤) نجد أنَّ $X$ متغيِّر عشوائي متقطِّع بمجموعة قيم غير منتهية ولكنَّها قابلة للعدّ.

٣- ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي ما $\left[\Omega,\mathscr{A},P\right]$ بحيث يمكن له أن يأخذ أيّة قيمة صحيحة موجبة، ومُعرَّف من خلال العلاقة الآتية:

$$P(X=n)=\frac{1}{n(n+1)} \qquad ;n\in\mathbb{N}$$

فنجد أنَّ:

$$\sum_{i\in I} p_i=\sum_{n=1}^{\infty} P(X=n)=\sum_{n=1}^{\infty}\frac{1}{n(n+1)}=1$$

وهنا نلاحظ أنَّ مجموعة قيم المتغيِّر العشوائي هي $\mathbf{X}=\mathbb{N}$، وهي مجموعة غير منتهية ولكنَّها قابلة للعدّ. إذاً فهو متغيِّر عشوائي متقطِّع.

**(٥,٣,١,٤) التمثيل الجدولي والبياني للمتغيِّرات العشوائيّة المتقطِّعة Tabular and graphic Representation of D.R.V.**

إذا كان $X$ متغيِّراً عشوائياً متقطِّعاً فوق فضاء احتمالي $\left[\Omega,\mathscr{A},P\right]$ بمجموعة قيم $\mathbf{X}=\{x_i\ ;i\in I\}$ فعندئذ من أجل أيّة قيمة $i\in I$ يمكننا أن نكتب الآتي:

$$P_X\left(\{x_i\}\right)=P\left(X^{-1}\left(\{x_i\}\right)\right)=P\left(X=x_i\right)=p_i$$

والتي تمثِّل قيمة الكتلية الاحتمالية للمتغيِّر العشوائي $X$ عند القيمة $x_i$، وفي الحالة الخاصَّة عندما تكون مجموعة قيم متغيِّر عشوائي متقطِّع $X$ منتهية (أي إنه متغيِّر عشوائي بسيط)، فإنَّه انطلاقاً من مُعرَّفة القيم $x_i$ و $P\left(X=x_i\right)$ يمكننا تقديم المتغيِّر العشوائي البسيط $X$ (ولو نظرياً على الأقل) وفق طرائق أخرى منها على سبيل المثال:

**١- التمثيل الجدولي**

بفرض أنَّ $X$ متغيِّر عشوائي بسيط مجموعة قيمه $\mathbf{X}=\left\{x_i\ ;i\in\mathbb{N}_n\right\}$ مع $n\in\mathbb{N}$ مثبَّت، فعندئذ يمكن تمثيل هذا المتغيِّر العشوائي جدولياً من خلال بناء جدول بصفين يدوَّن في صفه العلوي قيم $X$ وأمّا في صفه السفلي تدوَّن الاحتمالات الفردية المقابلة

لتلك القيم على النحو المقدَّم في الجدول الآتي:

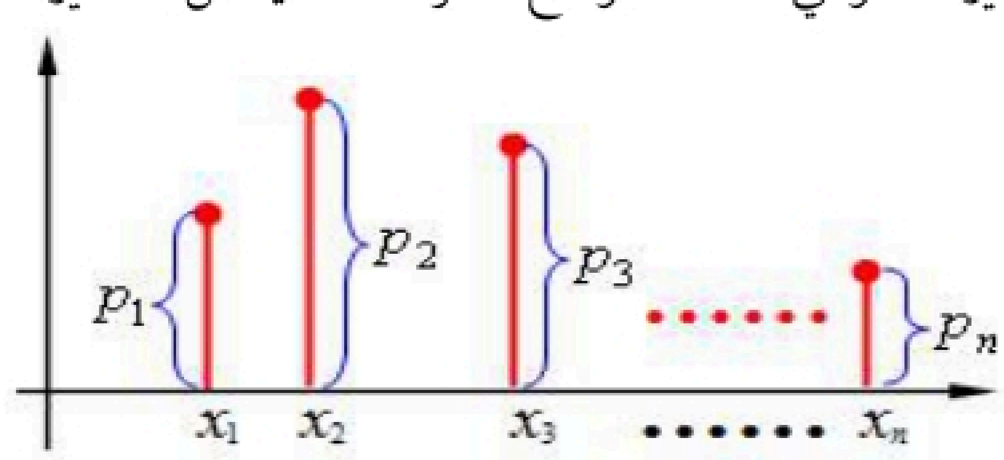
الجدول (٥,١)

| $i$ | 1 | 2 | ...... | $n$ | sum |
|---|---|---|---|---|---|
| قيم $X$ $x_i$ | $x_1$ | $x_2$ | ...... | $x_n$ | ... |
| $p_i = P(X = x_i)$ | $p_1$ | $p_2$ | ...... | $p_n$ | 1 |

ويُعرف هذا الجدول باسم **جدول توزيع المتغيِّر العشوائي $X$** أيضاً، وهذه التسميّة خاصّة بالمتغيِّرات العشوائيّة المتقطِّعة فقط.

**٢- التمثيل البياني**

بفرض أنَّ $X$ متغيِّر عشوائي بسيط مجموعة قيمه $\mathbf{X}=\{x_i \; ; i \in \mathbb{N}_n\}$ مع $n \ni \mathbb{N}$ مثبّت، فعندئذ يمكن تمثيل هذا المتغيِّر العشوائي بيانياً من خلال رسم محورين إحداثيين متعامدين، ومن ثمَّ رسم عمود بارتفاع يساوي $p_i$ (يُدعى قيمة قفزة للمتغيِّر العشوائي) فوق القيمة $x_i$ (وتُدعى موضع قفزة للمتغيِّر العشوائي) من أجل أيّة قيمة $i \ni \mathbb{N}_n$، ويوضع أحياناً في نهاية كل عمود دائرة صغيرة مُصْمَتة (وكأنَّها وزن) كتعبير عن موضع الكتلة الاحتمالية للمتغيِّر العشوائي عند موضع القفزة، وأمّا ارتفاع العمود (وهو الأهمّ في الرسم) فإنَّه يعبِّر عن قيمة الكتلة الاحتمالية للمتغيِّر العشوائي $X$ عند موضع القفزة. هكذا يمكن للمتغيِّر العشوائي $X$ أن يُمثَّل بيانياً من خلال الشكل الجانبي.




الشكل (٥,١٣)

**(٥,٣,١,٥) ملاحظات**

١- في الأشكال القادمة الممثِّلة للمتغيِّرات العشوائيّة المتقطِّعة سوف نتجاوز عن رسم الدوائر الصغيرة المُصْمَته في نهايات الأعمدة؛ وذلك لأنَّ المهمّ في هذه العروض هو ارتفاعات الأعمدة عند مواضع القفزات.

٢- يُلاحظ عدم جدوى تقديم المتغيِّرات العشوائيّة البسيطة بالطريقة الجدولية أو البيانية عندما تكون مجموعة قيمها كبيرة جداً، وأمّا إذا كانت القيم التي تأخذها هذه المتغيِّرات العشوائيّة قليلة فإنَّ هاتين الطريقتين تُعطيان عروضاً مقبولةً.

المبرهنة الآتية تقدّم لنا صيغة دالَّة التوزيع لمتغيِّر عشوائي متقطِّع.

**(٥,٣,١,٦) مبرهنة**

ليكن $X$ متغيِّراً عشوائياً متقطِّعاً فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$ بمجموعة قيم $\mathbf{X}=\{x_i \; ; i \in I\}$، فعندئذ من أجل كل $x \ni \mathbb{R}$ يكون لدالَّة توزيع $X$ العرض الآتي:

$$F_X(x) = \sum_{i \in I ; x_i < x} P(X = x_i) \qquad [5,17]$$

**البرهان:** ليكن $x \ni \mathbb{R}$ ولنفرض أنَّ $\mathbf{S}$ هي مجموعة الأدلة لقيم $X$ التي أصغر من $x$، فعندئذ يكون $x = \max\{x_i \; ; i \in \mathbf{S}\}$، ومن ثمَّ الاثبات يتم على النحو الآتي:

$$F_X(x) = P(X < x) = P\left(\{\omega \in \Omega \ ; X(\omega) < x_1\} \cup \{\omega \in \Omega \ ; X(\omega) < x_2\} \cup \ldots \cup \{\omega \in \Omega \ ; X(\omega) < x_{|s|}\}\right)$$
$$= P\left(X \in \sum_{i \in \mathbf{S}} \{x_i\}\right) = \sum_{i \in \mathbf{S}} P(X = x_i) = \sum_{i \in I ; x_i < x} P(X = x_i)$$

بهذا يتمّ البرهان.

**(٥,٣,١,٧) ملاحظات**

١- تجدر الإشارة هنا إلى أنَّ التوزيعات الاحتمالية النّاتجة عن متغيِّرات عشوائية متقطِّعة تُدعى **توزيعات متقطِّعة**.

٢- من العلاقة السّابقة [5,17] يُلاحظ أنَّ الرسم البياني لدوال التوزيع الاحتمالية للمتغيّرات العشوائيّة المتقطّعة هي **دوال درجية** دوماً (وقد لاحظنا ذلك في أمثلة سابقة).

**(٥,٣,١,٨) أمثلة**

١- ليكن لدينا $X$ متغيِّراً عشوائياً مقدَّماً جدولياً من خلال الجدول الآتي:

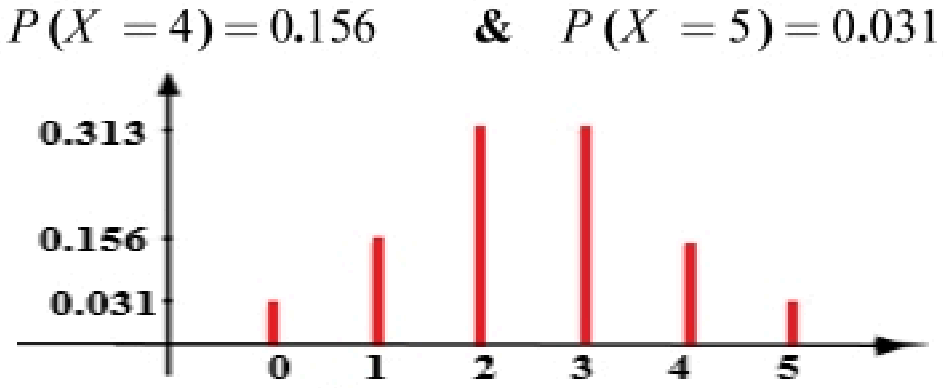
الجدول (٥,٢)

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | sum |
|---|---|---|---|---|---|---|---|
| قيم $X$ $x_i$ | 0 | 1 | 2 | 3 | 4 | 5 | ... |
| $P(X = x_i)$ | 0.031 | 0.156 | 0.313 | 0.313 | 0.156 | 0.031 | 1 |

فعندئذ نجد أنَّ هذا المتغيّر العشوائي متقطّع بسيط وتميّزه العلاقات الآتية:

$$P(X = 0) = 0.031 \quad \& \quad P(X = 1) = 0.165 \quad \& \quad P(X = 2) = 0.313$$
$$P(X = 3) = 0.313 \quad \& \quad P(X = 4) = 0.156 \quad \& \quad P(X = 5) = 0.031$$

ولتمثيله البياني الشـكل الآتي:



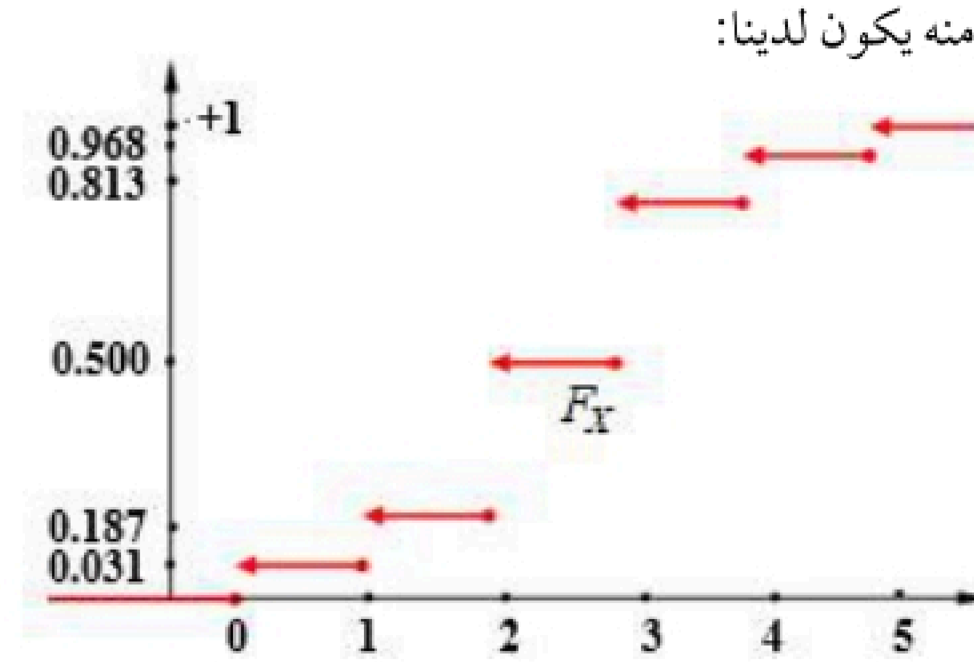

الشكل (٥,١٤.أ)

وأمّا لتعيين دالّة توزيعه الاحتمالية فلدينا من أجل كل $x \in \mathbb{R}$ ما يلي:

$$F_X(x) = \sum_{k=0 ; k < x}^{5} P(X = k)$$

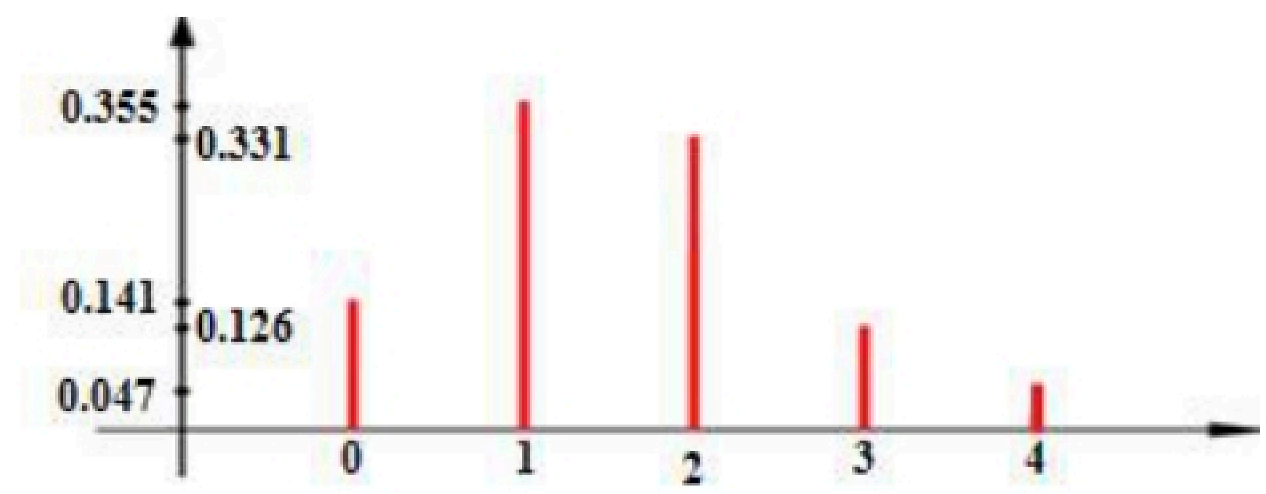
ومنه يكون لدينا:

$$F_X(x) = \begin{cases} 0 & \text{for } x \le 0 \\ 0.031 & \text{for } 0 < x \le 1 \\ 0.187 & \text{for } 1 < x \le 2 \\ 0.500 & \text{for } 2 < x \le 3 \\ 0.813 & \text{for } 3 < x \le 4 \\ 0.969 & \text{for } 4 < x \le 5 \\ 1 & \text{for } x > 5 \end{cases}$$




الشكل (٥,١٤.ب) الرسم البياني لدالّة التوزيع الاحتمالية.

٢- ليكن لدينا $X$ متغيِّراً عشوائياً مقدَّماً بيانياً من خلال الشكل الآتي:




الشكل (٥,١٥).

فعندئذ نجد لتمثيله الجدولي العرض الجدولي الآتي:

الجدول (٥,٣)

| $i$ | 1 | 2 | 3 | 4 | 5 | sum |
|---|---|---|---|---|---|---|
| قيم $X$ $x_i$ | 0 | 1 | 2 | 3 | 4 | ... |
| $P(X=x_i)$ | 0.141 | 0.355 | 0.331 | 0.126 | 0.047 | 1 |

ومن ثَمَّ العلاقات التي تميِّز هذا المتغيِّر العشوائي المتقطِّع البسيط هي:

$P(X=0)=0.141 \quad \& \quad P(X=1)=0.355$

$P(X=2)=0.126 \quad \& \quad P(X=3)=0.126$

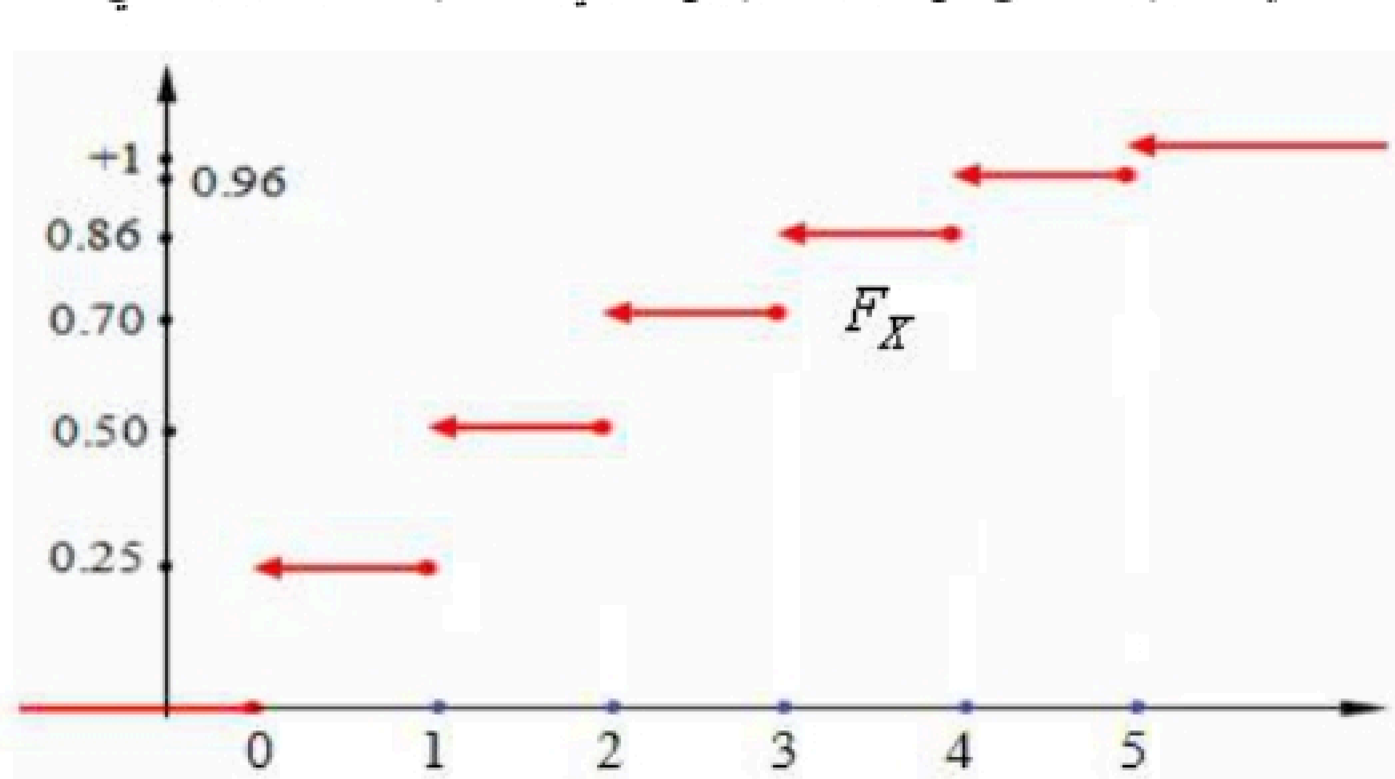
$P(X=4)=0.047$

وبحسابات بسيطة نجد العلاقة الآتية من أجل دالَّة التوزيع الاحتمالية لهذا المتغيِّر العشوائي:

$$F_X(x)=0.141\,\mathbf{I}_{(0,1]}(x)+0.496\,\mathbf{I}_{(1,2]}(x)+0.827\,\mathbf{I}_{(2,3]}(x)+0.953\,\mathbf{I}_{(3,4]}(x)+\mathbf{I}_{(4,+\infty)}(x)$$

وأمَّا الشكل البياني لهذه الدالَّة الدرجية فيترك للقارئ (يشابه الشكل (٥,١٤ ب)).

٣- ليكن لدينا $X$ متغيِّراً عشوائياً مُعطى من خلال العرض البياني لدالَّة توزيعه الاحتمالية التي لها الشكل الآتي:




الشكل (٥,١٦ أ)

فنجد أنَّ العرض التحليلي لدالَّة توزيع هذا المتغيِّر العشوائي هو:

$$F_X(x)=\begin{cases} 0 & for\ x\le 0\\ 0.25 & for\ 0<x\le 1\\ 0.50 & for\ 1<x\le 2\\ 0.70 & for\ 2<x\le 3\\ 0.86 & for\ 3<x\le 4\\ 0.96 & for\ 4<x\le 5\\ 1 & for\ x>5\end{cases}$$

ومن ثمَّ يكون للعلاقات التي تُميِّز هذا المتغيِّر العشوائي العروض الآتية:

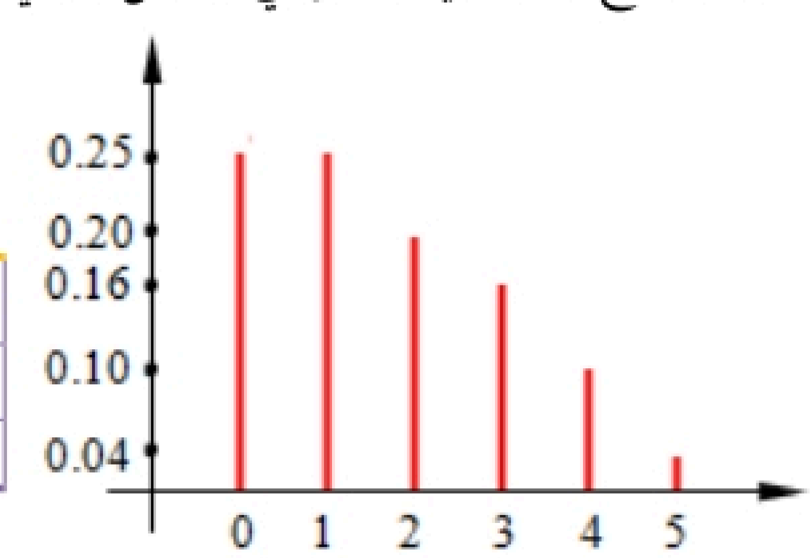
$$P(X=0)=0.25 \quad \& \quad P(X=1)=0.50-0.25=0.25 \quad \& \quad P(X=2)=0.70-0.50=0.20$$

$$P(X=3)=0.86-0.70=0.16 \quad \& \quad P(X=4)=96-0.86=0.10 \quad \& \quad P(X=5)=1-0.96=0.04$$

وكذلك يصبح لهذا المتغيِّر العشوائي التمثيل البياني والجدولي على النحو الآتي:

الجدول (٤ ,٥)

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | sum |
|---|---|---|---|---|---|---|---|
| قيم $X$ $x_i$ | 0 | 1 | 2 | 3 | 4 | 5 | ... |
| $P(X=x_i)$ | 0.25 | 0.25 | 0.20 | 0.16 | 0.10 | 0.04 | 1 |



الشكل (١٦ ,٥ ب). التمثيل لبياني لــ $X$

٤- لتكن لدينا الدالَّة الآتية:

$$f(x)=c(x^2+2) \qquad ; x=0,1,2,3,4$$

ولنقم بتعيين قيمة الثابت $c$ بحيث تصبح هذه الدالَّة هي دالَّة احتمالية لمتغيِّر عشوائي متقطِّع $X$.

من أجل ذلك نعلم أنَّ $\sum_{i\in I} P(X=x_i)=1$، ولذلك حتى تصبح الدالَّة $f(x)$ هي دالَّة احتمالية لمتغيِّر عشوائي متقطِّع $X$ يجب أن يكون $\sum_{x=0}^{4} f(x)=1$ (أي الدالَّة $f(x)$ أخذت دور $P(X=x_i)$)، ومنه يكون لدينا:

$$1=\sum_{x=0}^{4} f(x)=\sum_{x=0}^{4} c(x^2+2)=c(0^2+2)+c(1^2+2)+c(2^2+2)+c(3^2+2)+c(4^2+2)$$

$$=2c+3c+6c+11c+18c=40c$$

وهكذا نجد أنَّ $c=\frac{1}{40}$، ومن ثمَّ يكون للدالَّة الكتلية للمتغيِّر العشوائي $X$ العرض الآتي:

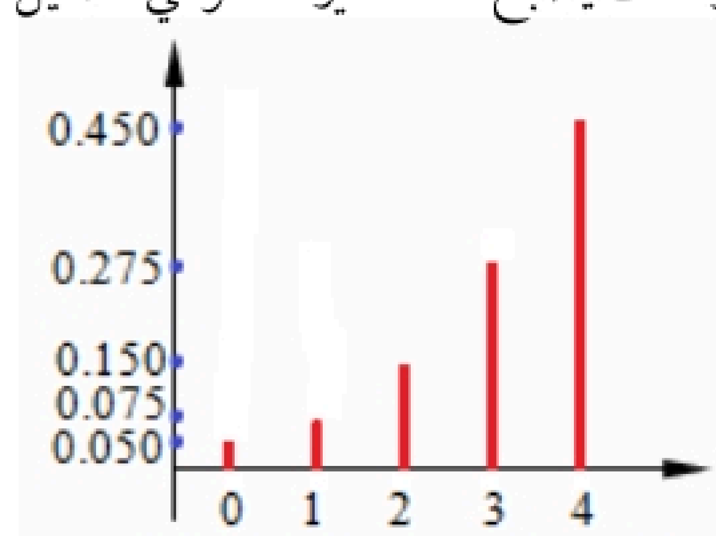
$$P(X=x)=\frac{x^2+2}{40} \qquad ; x=0,1,2,3,4$$

والتي نجد منها:

$$P(X=0)=0.050 \quad \& \quad P(X=1)=0.075 \quad \& \quad P(X=2)=0.150$$

$$P(X=3)=0.275 \quad \& \quad P(X=4)=0.450$$

وكذلك يصبح لهذا المتغيِّر العشوائي التمثيل الجدولي الآتي:




الشكل (١٧ , ٥). التمثيل لبياني لـ $X$.

الجدول (٥ , ٥)

| $i$ | 1 | 2 | 3 | 4 | 5 | sum |
|---|---|---|---|---|---|---|
| قيم $X$ $x_i$ | 0 | 1 | 2 | 3 | 4 | ... |
| $P(X = x_i)$ | 0.050 | 0.075 | 0.150 | 0.275 | 0.450 | 1 |

وبحسابات بسيطة نجد أنَّ لدالَّة التوزيع الاحتمالية لهذا المتغيِّر العشوائي العرض التحليلي الآتي:

$$F_X(x) = 0.05\ \mathbf{I}_{(0,1]}(x) + 0.125\ \mathbf{I}_{(1,2]}(x) + 0.275\ \mathbf{I}_{(2,3]}(x) + 0.55\ \mathbf{I}_{(3,4]}(x) + \mathbf{I}_{(4,+\infty)}(x)$$

**الخلاصة**: نلاحظ أنَّه من أجل متغيِّر عشوائي متقطِّع بسيط وبقيم قليلة العدد يمكننا استخدام أيّ من التمثيلات التي ذكرناها سابقاً لاستنتاج بقية التمثيلات الأخرى لهذا المتغيِّر العشوائي.

نكتفي بهذا القدر من المعلومات حول المتغيِّرات العشوائيّة المتقطِّعة حيث سنعود إليها مرَّة أخرى في فقرات وفصول قادمة، وننتقل فيما يلي إلى النوع الثاني من المتغيِّرات العشوائي.

### (٢ , ٣ , ٥) المتغيِّرات العشوائيّة المستمرَّة Continuous Random Variables

لقد ذكرنا سابقاً أنَّ المتغيِّرات العشوائيّة المستمرّة تتميَّز بدوال توزيع احتمالية مستمرَّة على $\mathbb{R}$، ولكنَّ التعريف الدقيق لهذا النوع من المتغيِّرات العشوائيّة يقدِّمه التعريف الآتي.

#### (١ , ٢ , ٣ , ٥) تعريف (المتغيّر العشوائي المستمرّ)

ليكن $X$ متغيِّراً عشـوائياً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، فإذا وجدت دالَّة حقيقية $f$ غير سالبة مُعرَّفة وقابلة للمكاملة على $\mathbb{R}$ وتُحقِّق العلاقة $\int_{-\infty}^{+\infty} f(x)\,dx = 1$، فعندئذ يُقال عن $X$ إنّه **مستمرّ** Continuous، وأمّا $f$ فإنَّها تُدعى **دالَّة الكثافة الاحتمالية لـ $X$**.

#### (٢ , ٢ , ٣ , ٥) ملاحظات

١- بناء على التعريف السّابق نلاحظ أنَّ المتغيِّرات العشوائيّة المستمرة تكون مُعيّنة (أو مُميَّزة) بوساطة دوال كثافتها الاحتمالية.

٢- إنَّ دالَّة الكثافة الاحتمالية $f$ المذكورة في التعريف السّابق هي في الواقع **دالَّة الكثافة لـ رادون- نيكوديم** Radon Nikodym Density Function من القياس $P_X$ بالنسبة إلى قياس لوبيغ $\lambda$. إنَّ اسم هذه الدالَّة يعود إلى الرياضياتي النمساوي **رادون** Johann Karl August Radon (1887–1956) والرياضياتي البولوني **نيكوديم** Otto Marcin Nikodym (1887–1974) (للاطلاع على بعض المفاهيم الواردة في هذه الفقرة ارجع إلى الملحق B)، ولتوضيح ذلك لنفترض أنَّه لدينا $X$ متغيِّر عشـوائي فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ وأنَّ قانون توزيعه $P_X$ مستمر مطلقاً بالنسبة إلى قياس لوبيغ $\lambda$ (علماً أنَّ $\lambda$ هو قياس من نوع $\sigma$- منته على الفضاء المقيس $[\mathbb{R}, \Re]$) ويعبَّر عن ذلك بالرمز $\lambda \prec\prec P_X$، فعندئذ الدالَّة $f$ تنتج من تفاضل رادون- نيكوديم للقياس $P_X$ بالنسـبة إلى قياس لوبيغ $\lambda$، أي إنَّ $f = \frac{d P_X}{d \lambda}$، وهي دالَّة غير سالبة وقيوسة بالنسبة إلى $\Re$ و $\Re \cap [0, \infty)$، ويُقال في هذه الحالة إنَّ المتغيِّر العشوائي $X$ **مستمر مطلقاً** Absolutely Continuous. من جهة أخرى، وبما أنَّ لكل متغيِّر عشوائي $X$ قانون توزيع خاص به $P_X$ فإنَّ كثافة رادون- نيكوديم النّاتجة عن هذا التفاضل ستكون خاصّة بالمتغيِّر العشوائي $X$، ولذلك يُرمز لها بـ $f_X$، أي إنَّه يُكتب:

$$\frac{d\,P_X}{d\,\lambda} = f_X \qquad [5,18]$$

ومن ثَّم يكون لدينا:

$$P_X\ (B) = \int_B f_X\,(x)\quad \lambda\,(dx) \qquad ;\ \forall B \in \Re \qquad [5,19]$$

أي إنَّ $f_X$ يجب أن تكون قابلة للمكاملة بحسب مفهوم لوبيغ، وفي حال كانت الدالَّة $f_X$ قابلة للمكاملة بحسب مفهوم ريمان فإنَّ العلاقة [5,19] يُكتب على سبيل التبسيط والاختصار كما يلي:

$$P_X\ (B) = \int_B f_X\,(x)\ \ dx \qquad ;\ B \in \Re \qquad [5,20]$$

وبما أنَّ $P_X$ قياس احتمالي على $[\mathbb{R}, \Re]$ فإنَّ العلاقة الآتية ستكون محقَّقة دوماً:

$$P_X\ (\mathbb{R}) = \int_{\mathbb{R}} f_X\,(x)\ \ dx \ \ = 1 \qquad [5,21]$$

علماً أنَّنا في دراستنا المقبلة لن نتعامل إلاّ مع دوال كثافة قابلة للمكاملة بحسب مفهوم ريمان.

٣- في الحقيقة يمكن تبيان (انظر [64]) أنَّه لكل دالَّة حقيقية $f$ معرّفة على $\mathbb{R}$ غير سالبة وقيوسة بالنسبة إلى $\Re$ و $\Re \cap [0\ ,\ \infty)$ ، وكمولةً (قابلة للمكاملة Integrable) بالنسبة إلى قياس لوبيغ $\lambda$ ، وتحقِّق العلاقة [5,21] يوجد قياس احتمالي وحيد $Q$ على $[\mathbb{R}, \Re]$ بحيث تكون $f$ هي كثافة رادون- نيكوديم من $Q$ بالنسبة إلى قياس لوبيغ $\lambda$ .

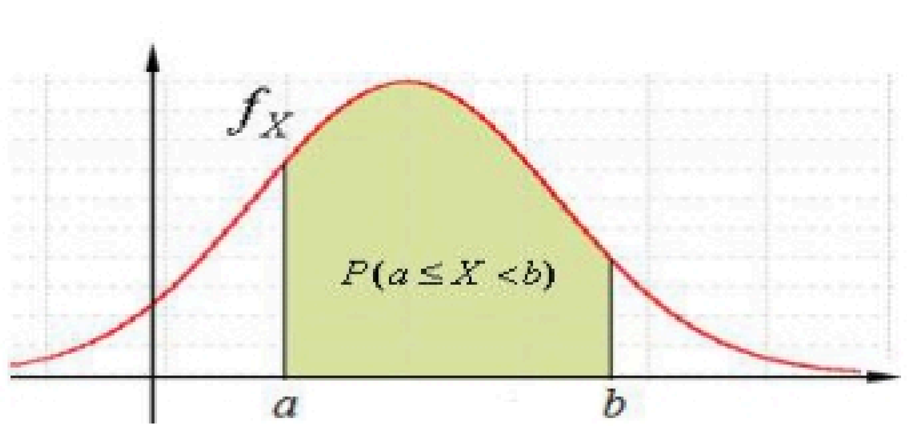
٤- يمكن تعريف المتغيِّر العشوائي المستمر (التعريف التقليدي) على النحو الآتي أيضاً:

يكون متغيِّر عشوائي $X$ فوق فضاء احتمالي $[\Omega, \mathcal{A}, P]$ مستمراً إذا وجدت من أجله دالَّة حقيقية غير سالبة $f_X$ مُعرَّفة وكمولة على $\mathbb{R}$ ، وبحيث إنَّه من أجل أي $a, b \in \mathbb{R}$ مع $b \geq a$ تكون العلاقة الآتية مُحقَّقة:

$$P\,(a \leq X\ \leq b\,) = \int_a^b f_X\ (x\,)\,dx \qquad [5,22]$$

ومن هذا التعريف للمتغيِّر العشوائي المستمر، وبحسب مفهوم ريمان لتكامل دالَّة نلاحظ أنَّ احتمال أن يأخذ متغيِّر عشوائي $X$ قيمة واقعة في فترة $[a\ ,\ b]$ يساوي للمساحة المحصورة بالمحور **oX** ومنحنى الدالَّة $f_X$ والمستقيمين $x = a$ و $x = b$ (انظر الشكل (٥, ١٨)). حيث يُلاحظ أنَّه عندما تسعى $b$ إلى $a$ فإنَّ المساحة المظللة التي تُعبِّر عن الاحتمال $P\,(a \leq X\ \leq b\,)$ تصغر أكثر فأكثر حتى إذا أصبحت $b = a$ فإنَّه يصبح لدينا $P\,(X\ = a) = 0$ ، وبما أنَّ $a$ قيمة كيفية من $\mathbb{R}$ فإنَّ سيكون لدينا:

$$P\,(X\ = a) = 0 \qquad ;\ \forall a \in \mathbb{R}$$



الشكل (٥, ١٨)

ومن ثَّم نحصل على النتيجة المهمَّة الآتية.

**(٥, ٣, ٢, ٣) نتيجة**

من أجل متغيِّر عشوائي مستمر $X$ تكون العلاقات الآتية مُحقَّقة دوماً:

$$P\,(a \leq X\ \leq b\,) = P\,(a < X\ \leq b\,) = P\,(a \leq X\ < b\,) = P\,(a < X\ < b\,)$$

توضيح آخر لهذه النتيجة تجدها في المثال القادم / ٢/ من (٥, ٣, ٢, ٧).

المبرهنة الآتية تعطينا صيغة دالَّة التوزيع لمتغيِّر عشوائي مستمر.

**(٥,٣,٢,٤) مبرهنة**

ليكن $X$ متغيِّراً عشوائياً مستمرّاً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ بدالَّة كثافة احتمالية $f_X$ ، فعندئذ من أجل كل $x \in \mathbb{R}$ يكون لدالَّة توزيع $X$ العرض الآتي:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\ dt \quad ; \forall x \in \mathbb{R} \qquad [5,23]$$

**البرهان:** لتكن $x \in \mathbb{R}$ قيمة كيفية، فعندئذ من تعريف وخواص دالَّة التوزيع الاحتمالية لمتغيِّر عشوائي، واستخدام العلاقة [5,23] يُمكننا أن نكتب الآتي:

$$F_X(x) = P(X < x) = P(X < x) - \lim_{a\to-\infty} P(X < a)$$

$$= \lim_{a\to-\infty} P(a \le X < x) = \lim_{a\to-\infty} P(a \le X \le x) = \lim_{a\to-\infty} \int_{a}^{x} f_X(t)\,dt = \int_{-\infty}^{x} f_X(t)\,dt$$

وبهذا يتمّ البرهان.

المبرهنة الآتية التي سنقدِّمها دون برهان (ذلك أنَّ برهانها يُقدَّم في مراحل متقدمة من نظرية الاحتمالات) توضِّح لنا الشروط الواجب توفرها في دالَّة حقيقية $f$ حتى تصبح دالَّة كثافة احتمالية لمتغيِّر عشوائي.

**(٥,٣,٢,٥) مبرهنة**

لتكن $f$ دالَّة حقيقية غير سالبة مُعرَّفة وكمولة على $\mathbb{R}$ ومُحقِّقة للعلاقة $\int_{-\infty}^{\infty} f(x)\ dx = 1$ ، فعندئذ يوجد متغيِّر عشوائي $X$ بحيث إنَّه من أجل كل $x \in \mathbb{R}$ يكون لدينا $f(x) = f_X(x)$.

**(٥,٣,٢,٦) نتائج**

١- من المبرهنة السّابقة وخواص دالَّة توزيع متغيِّر عشوائي نستنتج أنَّ:

$$\int_{-\infty}^{\infty} f_X(t)\,dt = 1 \qquad [5,24]$$

وذلك لأنَّه لدينا من خصائص دالَّة التوزيع الاحتمالية $\lim_{x\to\infty} F_X(x) = 1$ ، ومن ثَمَّ يمكننا أن نكتب الآتي:

$$\lim_{x\to\infty} F_X(x) = \lim_{x\to\infty} \int_{-\infty}^{x} f_X(t)\,dt = \int_{-\infty}^{\infty} f_X(t)\,dt = 1$$

٢- ليكن $X$ متغيِّراً عشوائياً مستمرّاً بدالَّة كثافة احتمالية $f_X$ مستمرَّة في نقطة $x \in \mathbb{R}$ ، فعندئذ (بحسب مبرهنة في حساب التفاضل) تكون دالَّة توزيعه الاحتمالية $F_X$ قابلة للمفاضلة في هذه النقطة $x$، ولدينا:

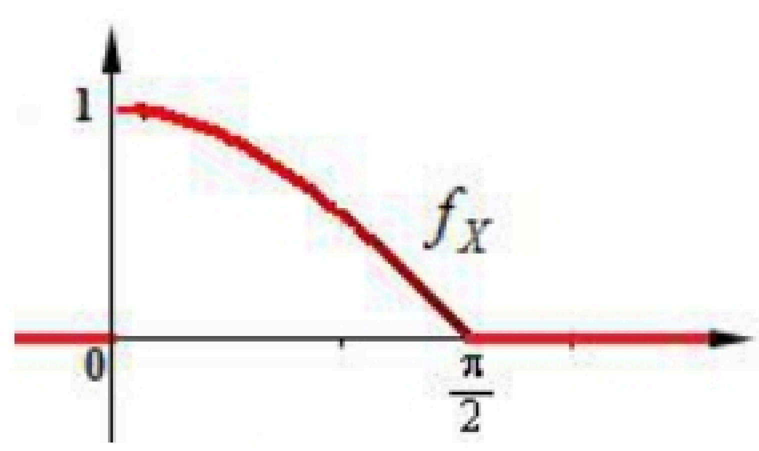
$$\frac{d\,F_X(x)}{dx} = f_X(x) \qquad [5,25]$$

ومن ثَمَّ دالَّة توزيع متغيِّر عشوائي مستمر $X$ تكون مستمرَّة على $\mathbb{R}$ ؛ وذلك لأنَّها بمثابة الدالَّة الأصلية لدالَّة حقيقية $f_X$ كمولة على $\mathbb{R}$ (أي لـ $f_X$ على الأكثر عدد منته من نقاط الانقطاع من النوع الأول). وعندما تكون الدالَّة $f_X$ ليست مستمرَّة في نقطة $x \in \mathbb{R}$ ، فعندئذ ستكون الدالَّة $F_X$ غير قابلة للمفاضلة في هذه النقطة $x$، وهذا يعني أنَّ الرسم البياني للدالَّة $F_X$ سوف يعاني من انكسار أو تدبّب عند هذه النقطة، والأمثلة الآتية توضِّح لنا ذلك.

**(٧, ٢, ٣, ٥) أمثلة**

١ - ليكن $X$ متغيِّراً عشوائياً مستمرَّاً بدالَّة كثافة احتمالية:

$$f_X(x) = \cos(x) \cdot \mathbf{I}_{[0\,,\,\pi/2]}(x) \quad ; \forall x \in \mathbb{R}$$




الشكل (١٩ , ٥. أ) الرسم البياني لدالَّة الكثافة الاحتمالية

فنلاحظ أنَّ دالَّة الكثافة الاحتمالية لهذا المتغيِّر العشوائي تعاني من انقطاع عند النقطة $x = 0$، ولنقم بتعيين دالَّة توزيع $X$.

الحـل: لتعيين دالَّة توزيع $X$ سنناقش الحالات الآتية:

أ- من أجل $x \leq 0$ لدينا الآتي:

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt \underset{for\ x\leq 0}{=\!=\!=\!=\!=} \int_{-\infty}^{x} \overbrace{f_X(t)}^{0}\,dt = 0$$

ب- من أجل $0 < x \leq \pi/2$ لدينا الآتي:

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt \underset{for\ 0<x\leq\frac{\pi}{2}}{=\!=\!=\!=\!=\!=\!=} \int_{-\infty}^{0} \overbrace{f_X(t)}^{0}\,dt + \int_{0}^{x} \overbrace{f_X(t)}^{\cos(t)\cdot\mathbf{I}_{[0,\pi/2]}(t)}\,dt$$

والتي نجد منها:

$$F_X(x) = \int_{0}^{x} \cos(t)\cdot\mathbf{I}_{[0\,,\,\pi/2]}(t)\,dt = \int_{0}^{x} \cos(t)\,dt = \left(\sin(t)\Big|_{0}^{x}\right) = (\sin(x) - 0) = \sin(x)$$

ج- من أجل $x > \frac{\pi}{2}$ لدينا الآتي:

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt \underset{for\ x>\frac{\pi}{2}}{=\!=\!=\!=\!=} \int_{-\infty}^{0} \overbrace{f_X(t)}^{0}\,dt + \int_{0}^{\frac{\pi}{2}} \overbrace{f_X(t)}^{\cos(t)\cdot\mathbf{I}_{[0,\pi/2]}(t)}\,dt + \int_{\frac{\pi}{2}}^{x} \overbrace{f_X(x)}^{0}\,dx$$

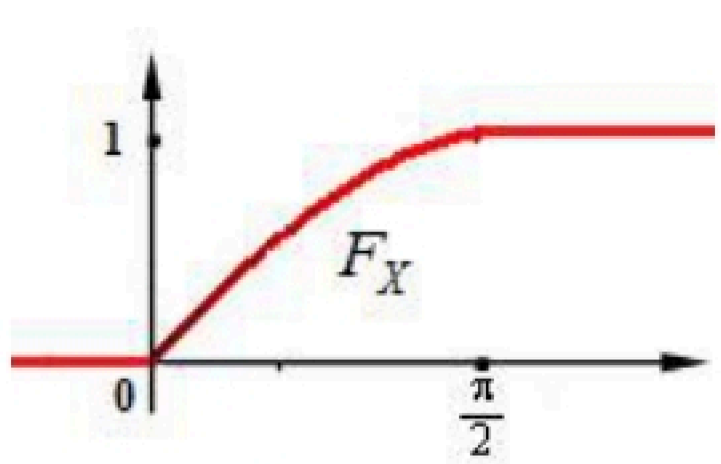
$$= \int_{0}^{\frac{\pi}{2}} \cos(t)\,dt = \left(\sin t\Big|_{0}^{\frac{\pi}{2}}\right) = (1-0) = 1$$

ومن ثمَّ يكون لدالَّة التوزيع الاحتمالية $F_X$ العرض الآتي:

$$F_X(x) = \sin(x)\cdot\mathbf{I}_{[0\,,\,\pi/2)}(x) + \mathbf{I}_{[0\,,\,\infty)}(x) \quad ; \forall x \in \mathbb{R}$$

والتي يمكن عرضها وفقاً للصيغة الشرطية الآتية أيضاً:

$$F_X(x) = \begin{cases} 0 & for\ x \leq 0 \\ \sin(x) & for\ 0 < x \leq \pi/2 \\ 1 & for\ x > \pi/2 \end{cases}$$




الشكل (١٩ , ٥. ب) الرسم البياني لدالَّة التوزيع الاحتمالية

حيث نلاحظ أنَّ دالَّة التوزيع الاحتمالية لهذا المتغيِّر العشوائي غير قابلة للمفاضلة عند النقطة $x = 0$.

٢- ليكن $X$ متغيِّراً عشوائياً مستمرّاً بدالّة كثافة احتمالية $f_X$ مُعطاة من خلال العلاقة الآتية:

$$f_X(x) = c \cdot \mathbf{I}_{[0,10]}(x) \qquad ; \forall x \in \mathbb{R}$$

علماً أنَّ $c$ ثابت حقيقي، ولنقم بتعيين قيمة هذا الثابت ومن ثَمَّ دالّة التوزيع الاحتمالية $F_X$ وبعد ذلك تقديم الرسم البياني لكل من $f_X$ و $F_X$.

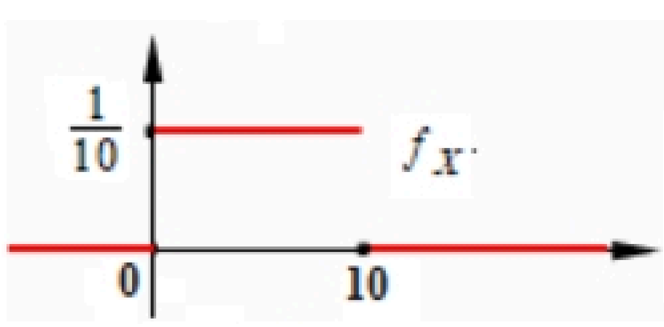
الحل: لتعيين الثابت $c$ نعلم أنَّ:

$$1 = \int_{-\infty}^{\infty} f_X(x)\ dx = c \int_{0}^{10} dx = c \cdot (10-0) = 10\,c$$

التي نجد منها أنَّ $c = \frac{1}{10}$، ومن ثَمَّ يصبح لدالّة الكثافة الاحتمالية $f_X$ العرض الآتي:

$$f_X(x) = \frac{1}{10}\,\mathbf{I}_{[0,1]}(x) \qquad ; \forall x \in \mathbb{R}$$

فنلاحظ أنَّ دالّة الكثافة الاحتمالية لهذا المتغيِّر العشوائي تعاني من نقطتي انقطاع عند النقطة $x = 0$ و $x = 10$، ولذلك سوف نلاحظ أنَّ دالّة التوزيع الاحتمالية لهذا المتغيِّر العشوائي ستكون غير قابلة للمفاضلة عند النقطتين $x = 0$ و $x = 10$.



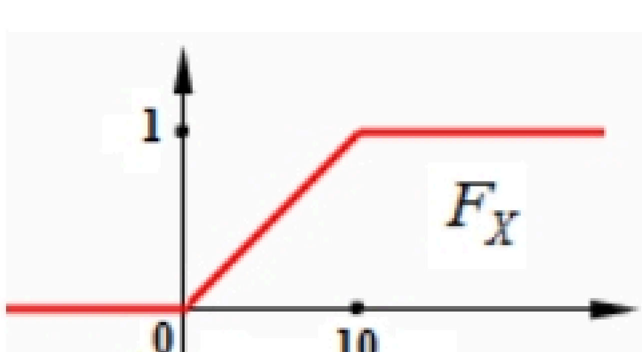
الشكل (٢٠, ٥. أ) الرسم البياني لدالّة الكثافة الاحتمالية

الآن لتعيين دالّة التوزيع الاحتمالية $F_X$ نجد وبأسلوب مماثل لحل المثال السّابق أنّه من أجل كل $x$ في $\mathbb{R}$ لدينا:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\ dt = \frac{1}{10}\int_{-\infty}^{x} \mathbf{I}_{[0,10]}(t)\ dt = \frac{x}{10}\cdot \mathbf{I}_{(0,10]}(x) + \mathbf{I}_{(10,\infty)}(x)$$

والتي يمكن كتابتها على النحو الآتي أيضاً:

$$F_X(x) = \begin{cases} 0 & for\ x \le 0 \\ \dfrac{x}{10} & for\ 0 < x \le 10 \\ 1 & for\ x > 10 \end{cases}$$



الشكل (٢٠, ٥. ب) الرسم البياني لدالّة التوزيع الاحتمالية

يمكننا باستخدام هذا المثال تقديم توضيح آخر للوصول إلى النتيجة التي تقول إنَّه من أجل $X$ متغيِّر عشوائي مستمر، و $a$ قيمة حقيقية ثابتة يكون لدينا $P(X = a) = 0$، وذلك على النحو الآتي:

نعلم من المبرهنة (٥, ٢, ٢, ٣) أنَّ $P(a \le X \le b) = F_X(b) - F_X(a)$، ومن ثَمَّ باستخدام المتغيّر العشوائي المُعطى في هذا المثال يكون لدينا ما يلي:

$$P(2 \le X \le 4) = F_X(4) - F_X(2) = 0.2$$

$$P(2.5 \le X \le 3.5) = F_X(3.5) - F_X(2.5) = 0.1$$

$$P(2.95 \le X \le 3.05) = F_X(3.05) - F_X(2.95) = 0.01$$

$$P(2.995 \le X \le 3.005) = F_X(3.005) - F_X(2.995) = 0.001$$

$$P(2.9995 \le X \le 3.0005) = F_X(3.0005) - F_X(2.9995) = 0.0001$$

وهكذا نلاحظ أنَّه عندما القيمة الراجحة على $X$ من الأعلى تسعى إلى الـ 3 وفي الوقت نفسه القيمة المرجوحة بـ $X$ من الأدنى تسعى إلى الـ 3 أيضاً، فإنَّ قيمة الاحتمال المقدَّم سوف يسعى إلى $P(X = 3)$ ، وأمّا القيمة المقابلة لهذا الاحتمال فإنَّها تسعى إلى الصفر وضوحاً.

٣- ليكن $X$ متغيِّراً عشوائياً مستمرّاً بدالَّة كثافة احتمالية $f_X$ مُعطاة من خلال العلاقة الآتية:

$$f_X(x)=\begin{cases} c\sqrt{x} & \text{for } 0<x<1 \\ 0 & \text{otherwise}\end{cases}$$

ولنقم بما يلي:

أ- تعيين قيمة الثابت $c$، ومن ثَّم رسم بيان الدالَّة $f_X$.

ب- تعيين دالَّة التوزيع الاحتمالية لـ $X$.

ج- استخدم دالَّة التوزيع الاحتمالية لـ $X$ في حساب الاحتمال $P(0.3 < X < 0.6)$.

**الحل**: من أجل الفقرة:

أ- نعلم أنَّ $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$ ، ومن ثَّم يكون لدينا:

$$1=\int_{-\infty}^{\infty} f_X(x)\,dx = \int_{-\infty}^{0}\overbrace{f_X(x)}^{0}dx+\int_{0}^{1}\overbrace{f_X(x)}^{c\sqrt{x}}dx+\int_{1}^{\infty}\overbrace{f_X(x)}^{0}dx$$

$$=\int_0^1 c\sqrt{x}\,dx = c\left(\frac{2}{3}x^{\frac{3}{2}}\Big|_0^1\right)=\frac{2c}{3}(1-0)=\frac{2c}{3}$$

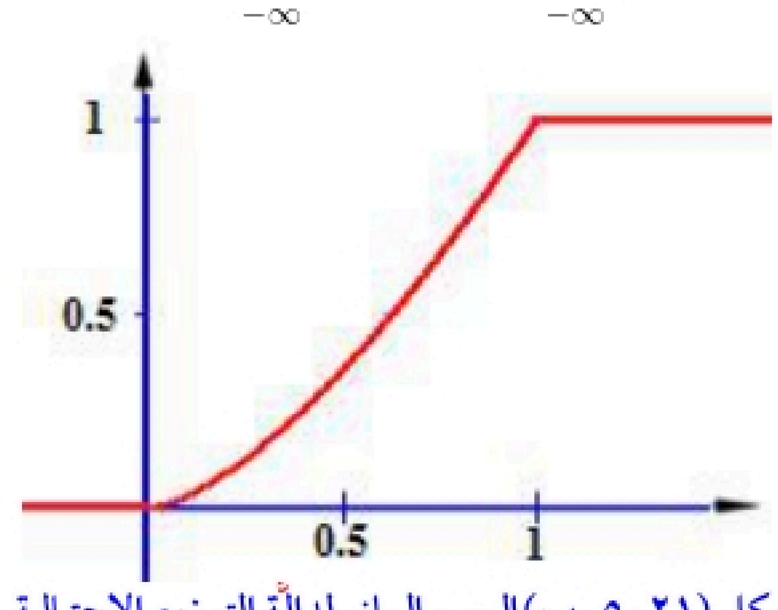
والتي نجد منها أنَّ $c=\frac{3}{2}$، ومن ثَّم يكون لدالَّة الكثافة الاحتمالية $f_X$ العرض الآتي.

$$f_X(x)=\begin{cases} \frac{3}{2}\sqrt{x} & \text{for } 0<x<1 \\ 0 & \text{eleswhere}\end{cases}$$

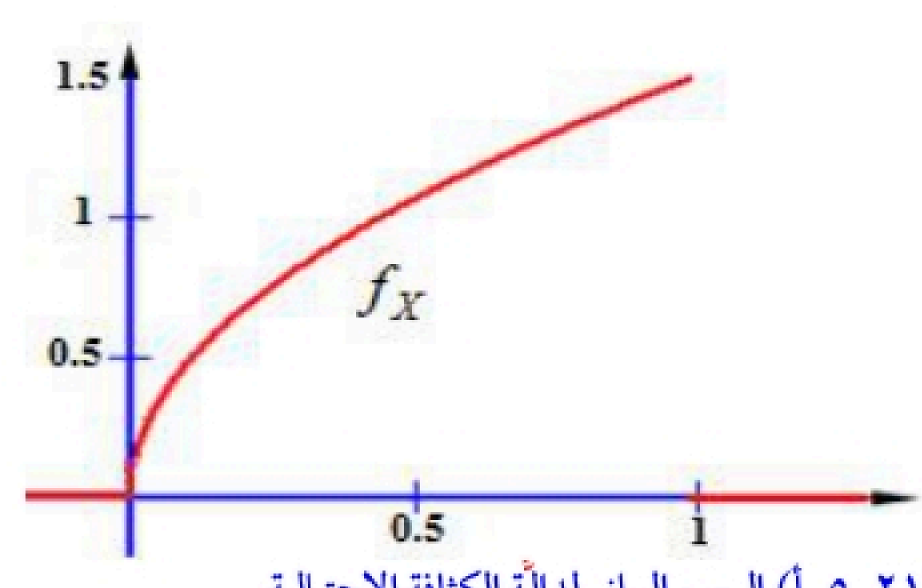
فنلاحظ أنَّ دالَّة الكثافة الاحتمالية لهذا المتغيِّر العشوائي تعاني من نقطة انقطاع عند $x = 1$ (انظر الشكل (٢١,٥. أ))، ولذلك سوف نلاحظ أنَّ دالَّة التوزيع الاحتمالية لهذا المتغيِّر العشوائي ستكون غير قابلة للمفاضلة عند هذه النقطة (انظر الشكل (٢١,٥. ب)).

ب- لتعيين دالَّة التوزيع الاحتمالية $F_X$ نجد وبأسلوب مماثل لحل المثال الأول السّابق أنَّه من أجل كل $x \in \mathbb{R}$ لدينا:

$$F_X(x)=\int_{-\infty}^{x} f_X(t)\,dt=\int_{-\infty}^{x}\frac{3}{2}t^{\frac{1}{2}}\cdot \mathbf{I}_{[0,1]}(t)\,dt = x^{\frac{3}{2}}\cdot \mathbf{I}_{(0,1]}(x)+\mathbf{I}_{(1,\infty)}(x)$$




الشكل (٢١,٥. ب) الرسم البياني لدالَّة التوزيع الاحتمالية




الشكل (٢١,٥. أ) الرسم البياني لدالَّة الكثافة الاحتمالية

ج- لحساب $P(0.3 < X < 0.6)$ لدينا:

$$P(0.3 < X < 0.6) = F_X(0.6) - F_X(0.3) = 0.6^{\frac{3}{2}} - 0.3^{\frac{3}{2}} = 0.465 - 0.164 = 0.301$$

وبهذا ينتهي حل المثال.

٤- لتكن $f$ دالَّة حقيقية مُعرَّفة على $\mathbb{R}$ كما يلي:

$$f(x) = \frac{1}{\mathbf{e}-1}\mathbf{e}^x \cdot \mathbf{I}_{[0,1]}(x) \quad ; \forall x \in \mathbb{R}$$

فنلاحظ أنَّ هذه الدالَّة غير سالبة على $\mathbb{R}$، ولدينا:

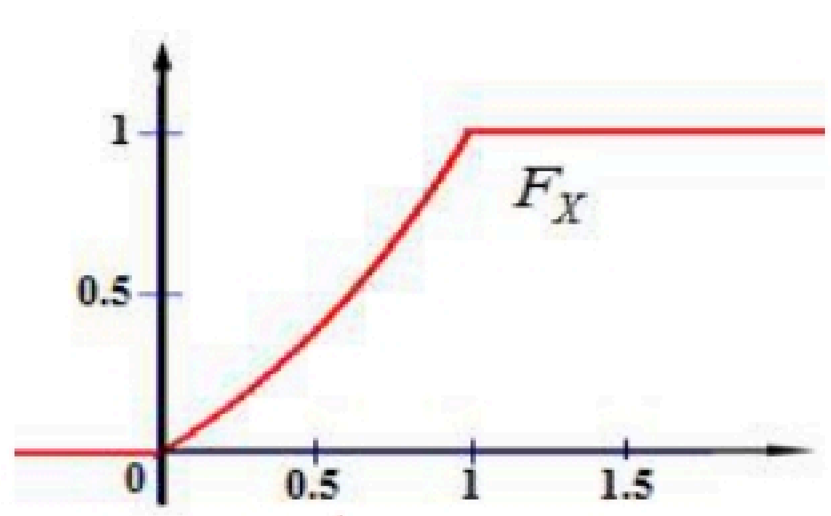
$$\int_{-\infty}^{\infty} f(x)\, dx = \frac{1}{\mathbf{e}-1}\int_{0}^{1}\mathbf{e}^x\, dx = 1$$

وهذا يعني أنَّ هذه الدالَّة تصلح لأن تكون دالَّة كثافة احتمالية لمتغيِّر عشوائي $X$، بمعنى أنَّه يوجد متغيِّر عشوائي $X$ بحيث تكون هذه الدالَّة هي دالَّة كثافة احتمالية له للاطلاع على رسمها البياني (انظر الشكل (٢٢ ,٥. أ)).

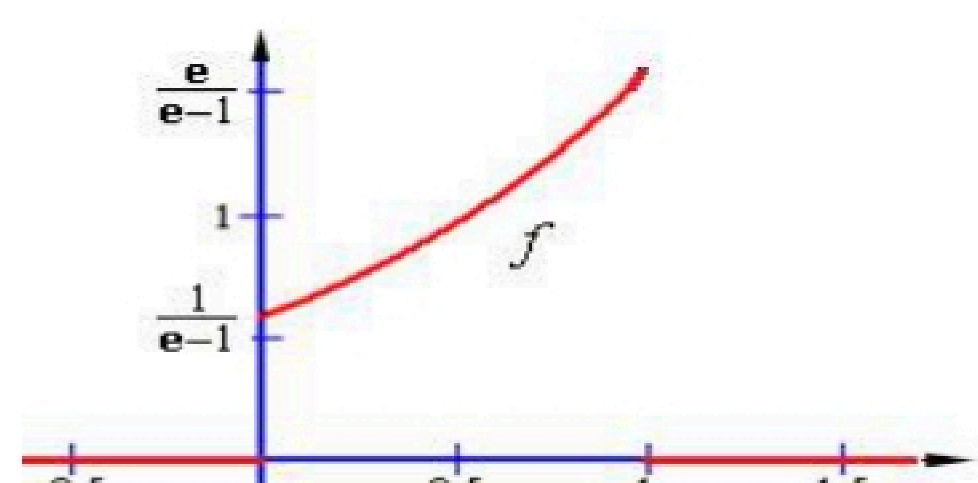
الآن لو رمزنا لهذه الدالَّة بـ $f_X$، فعندئذ يكون لدالَّة التوزيع $F_X$ العرض الآتي من أجل كل $x \in \mathbb{R}$:

$$F_X(x) = \int_{-\infty}^{\infty} f_X(x)\, dx = \frac{1}{\mathbf{e}-1}\int_{-\infty}^{x}\mathbf{e}^t\, dt = \frac{\mathbf{e}^x - 1}{\mathbf{e}-1}\mathbf{I}_{(0,1]}(x) + \mathbf{I}_{(1,\infty)}(x)$$

حيث نلاحظ بوضوح أنَّ دالَّة التوزيع الاحتمالية لـ $X$ هي دالَّة غير قابلة للمفاضلة عند النقطة $x = 1$ (انظر الشكل (٢٢ ,٥. ب)).




الشكل (٢٢ ,٥. ب) الرسم البياني لدالَّة التوزيع الاحتمالية




الشكل (٢٢ ,٥. أ) الرسم البياني لدالَّة الكثافة الاحتمالية

٥- لتكن $F_X$ دالَّة توزيع احتمالية لمتغيِّر عشوائي $X$ مُعطاة من خلال العلاقة الآتية (ورسمها البياني يقدِّمه الشكل(٢٣ ,٥. أ)):

$$F_X(x) = \left(4x^3 - 6x^2 + 3x\right)\mathbf{I}_{(0,1]}(x) + \mathbf{I}_{(1,\infty)}(x) \quad ; \forall x \in \mathbb{R}$$

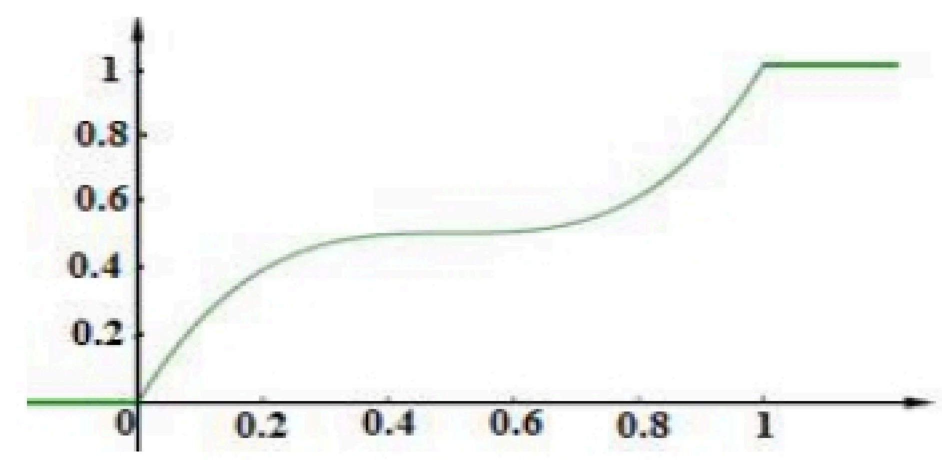
ولنقم بتعيين دالَّة الكثافة الاحتمالية الموافقة لـ $F_X$.

الحـل: باستخدام العلاقة [5,25] يكون من أجل كل $x \in (0,1)$ لدينا ما يلي:

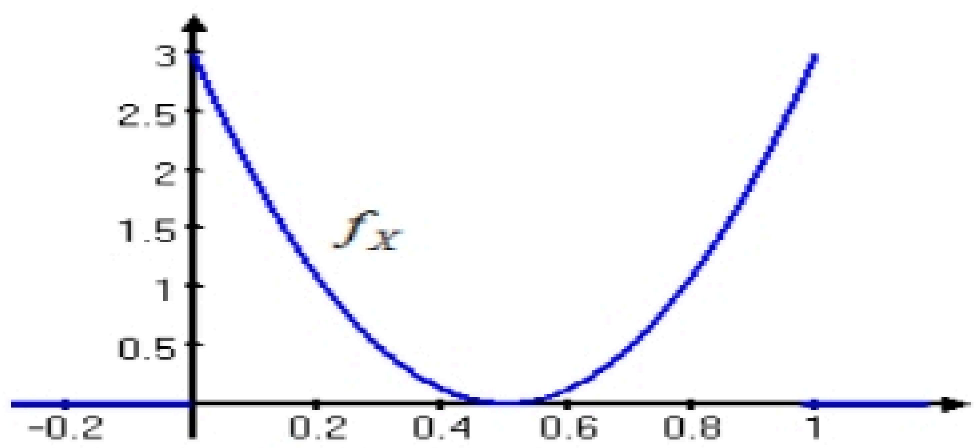
$$f_X(x) = \frac{d\,F_X(x)}{dx} = \left(12x^2 - 12x + 3\right)\mathbf{I}_{(0,1)}(x)$$

وأمَّا من أجل قيم $x$ التي لا تنتمي للفترة $(0,1)$ فلدينا $f_X(x) = 0$ لأنَّ $F_X$ ثابتة على $\mathbb{R}\setminus(0,1)$، وأخيراً الرسم البياني لـ $F_X$

نجده في الشكل (٢٣ , ٥ . ب) الآتي.




الشكل (٢٣ , ٥ . أ) الرسم البياني لدالَّة التوزيع الاحتمالية




الشكل (٢٣ , ٥ . ب) الرسم البياني لدالَّة الكثافة الاحتمالية

## (٥,٤) المتجهات العشوائية وتوزيعاتها الاحتمالية

**Random Vectors and their Probability Distributions**

لقد قمنا فيما سبق بدراسة متغيِّرات عشوائيَّة وحيدة البعد. فيما يلي سنقدِّم دراسة لمتغيِّرات عشوائيَّة متعدِّدة الأبعاد يُطلق عليها اسم متَّجهات عشوائيَّة Random Vectors. إنَّ ضرورة التعامل مع متغيِّرات عشوائيَّة متعدّدة الأبعاد ينبع من متطلبات الدراسة الاحتمالية لبعض التجارب العشوائيّة في مجالات التصنيع، والاتصالات، وعلم الحياة، وعلوم كثيرة أخرى، حيث يقف المتغيِّر العشوائي وحيد البعد عاجزاً عن تمثيل مثل تلك التجارب أو دراستها، فعلى سبيل المثال لو أردنا القيام بإجراء دراسة حول تطور الطول والوزن مع عمر الأطفال حتى سن مُعيِّن، فعندئذ النتائج المتمخضة عن هذه التجربة تعطينا ثلاثة نماذج من القيم العشوائيّة الأول خاص بالقيم النّاتجة عن الطول ونحتاج لذلك متغيِّراً عشوائياً $X_1$، والثاني خاص بالقيم النّاتجة عن الوزن ويمثِّله متغيِّر عشوائي ثان $X_2$، وأخيراً متغيِّر عشوائي ثالث خاص بالقيم النّاتجة عن العمر ويمثِّله متغيِّر عشوائي آخر $X_3$، وبالمحصلة نجد أنَّ نتائج دراسة كلِّ طفل يُمثَّل بثلاثية $(x_1, x_2, x_3)$ كقيمة عشوائيَّة للمتَّجه $(X_1, X_2, X_3)$ لدى تطبيقه على ذلك الطفل، ويُلاحظ هنا أنَّ المتَّجه $(X_1, X_2, X_3)$ يأخذ قيمة في $\mathbb{R}^3$. بالطبع كان بإمكاننا توسيع أبعاد هذا المتَّجه إذا ما قمنا بإضافة أي ظاهرة من ظواهر الجسم البشري مثل الحرارة ومحيط الرأس و...

إنَّ الخطوة الأولى لتعميم الدراسة السّابقة هي دراسة التوزيعات الاحتمالية على $\mathbb{R}^n$ مع $n \in \mathbb{N}$ التي ينتج عنها دراسة توزيعات المتَّجهات العشو ائيّة علماً أنّنا سنركز دراستنا على المتَّـجهات العشوائيّة الحقيقية فقط، وأكثر من ذلك، لن نتوسَّع كثيراً في دراسة هذا الموضوع بل سنكتفي ببعض المفاهيم والنتائج الأساسية والمهمّة.

### (١ , ٤ , ٥) مفهوم المتَّجه العشوائي

لقد قمنا سابقاً بتقديم المتَّجه العشوائي من خلال أحد الأمثلة على العناصر العشوائيَّة (المثال / ٢/ من الفقرة (١ , ٢ , ١ , ٥))، وفيما يلي سنعيد صياغته على نحو آخر كالآتي.

لتكن $X_1$، $X_2$، ... و $X_n$ متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، فعندئذ نجد أنَّ التطبيق:

$$(X_1, X_2, \ldots, X_n) : \Omega \longrightarrow \mathbb{R}^n$$

$$\omega \mapsto (X_1, X_2, \ldots, X_n)(\omega) := (X_1(\omega), X_2(\omega), \ldots, X_n(\omega))$$

هو عنصر عشوائي فوق $[\Omega, \mathscr{A}, P]$ وبقيم في الفضاء المقيس $[\mathbb{R}^n, \Re^n]$؛ وذلك لأنَّه من أجل أي $B_1$، $B_2$، ... و $B_n$ من $\Re$ لدينا ما يلي مُحقَّقاً:

$$\left(X_1, X_2, \dots, X_n\right)^{-1}\left(B_1 \times B_2 \times \dots \times B_n\right) = \{\omega \in \Omega\ ;\ X_1(\omega) \in B_1 \wedge X_2(\omega) \in B_2 \wedge \dots \wedge X_n(\omega) \in B_n\}$$
$$= \bigcap_{k=1}^{n} X_k^{-1}\left(B_k\right) \in \mathfrak{F}$$

وذلك لأنَّ $X_1$، $X_2$،... و $X_n$ متغيِّرات عشوائية فوق $[\Omega, \mathscr{A}, P]$ و $\mathscr{A}$ مغلق بالنسبة إلى عملية التقاطع المنتهي وغير المنتهي العدود. علماً أنَّ $\wedge$ هو الرمز المنطقي الذي يشير إلى عبارة ( و **and**)، وهذا الرمز يكثر استخدامه في مجال دراسة المتَّجهات العشوائيّة.

هكذا يمكننا إعاة تقديم المتَّجه العشوائي من خلال التعريف الآتية.

### (٥, ٤, ١, ١) تعريف (المتَّجه العشوائي)

لتكن $X_1$، $X_2$،... و $X_n$ متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، فعندئذ يُدعى $(X_1, X_2, \dots, X_n)$ **متَّجهاً عشوائياً** ذا بُعْد يساوي $n$ (أو منته البُعد)، وكنّا قد رمزنا له بـ $\mathbb{X}_n$، أي إنَّنا سنكتب:

$$\mathbb{X}_n \triangleq \left(X_1, X_2, \dots, X_n\right) \qquad [5,26]$$

### (٥, ٤, ١, ٢) ملاحظات

١- يمكننا من أجل أي متَّجه عشوائي $\mathbb{X}_n$ فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ إيجاد $n$ متغيِّر عشوائي $X_1$، $X_2$،... و $X_n$ فوق $[\Omega, \mathscr{A}, P]$ بحيث يكون من أجل كل $\omega \in \Omega$ لدينا:

$$\mathbb{X}_n(\omega) = \left(X_1(\omega), X_2(\omega), \dots, X_n(\omega)\right) \qquad [5,27]$$

ولإثبات ذلك لنمعن النظر في التركيب الآتي:

$$\Omega \xrightarrow{\quad \mathbb{X}_n \quad} \mathbb{R}^n \xrightarrow{\quad \Pi_i \quad} \mathbb{R}$$
$$\omega \mapsto \mathbb{X}_n(\omega) = \left(x_1, \dots, x_n\right) \mapsto \Pi_i\left(x_1, \dots, x_n\right) = x_i$$

علماً أنَّ $\Pi_i$ مع $i \in \mathbb{N}_n$ هي **دوال الإسقاط** Projection Functions على $\mathbb{R}^n$، فعندئذ لو أخذنا:

$$X_i = \Pi_i \circ \mathbb{X}_n \qquad ;\ i \in \mathbb{N}_n \qquad [5,28]$$

فإنَّنا نجد أنَّ $X_1$، $X_2$،... و $X_n$ هي متغيّرات عشوائية فوق $[\Omega, \mathscr{A}, P]$، وهذا يعني أنَّه من أجل أي متَّجه عشـوائي $\mathbb{X}_n$ على $[\Omega, \mathscr{A}, P]$ يوجد بالفعل $n$ متغيِّر عشوائي $X_1$، $X_2$،... و $X_n$ فوق $[\Omega, \mathscr{A}, P]$ بحيث تكون العلاقة [5,28] مُحقَّقةً.

٢- إذا كان $\mathbb{X}_n = \left(X_1, X_2, \dots, X_n\right)$ متَّجهاً عشوائياً مُركِّباته $X_1$، $X_2$،... و $X_n$ متغيِّرات عشوائية عقدية، فعندئذ يُقال عن $\mathbb{X}_n$ إنَّه **متَّجه عشوائي عقدي**.

### (٥, ٤, ٢) دالَّة التوزيع الاحتمالية لمتَّجه عشوائي Distribution Function of a Random Vector

الآن، وبما أنَّ قيم المتَّجه العشوائي $\mathbb{X}_n$ في $\mathbb{R}^n$ عشوائية، فإنَّه سيقابل الفضاء الاحتمالي $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً مُولَّداً من $\mathbb{X}_n$، وهذا يعني ضرورة وجود قياس احتمالي على $\left[\mathbb{R}^n, \mathfrak{R}^n\right]$ مولَّداً من $\mathbb{X}_n$، وهذا بدوره يوصلنا إلى مفهوم قانون توزيع متَّجه عشوائي والذي سنقدِّمه من خلال التعريف الآتي.

### (٥, ٤, ٢, ١) تعريف (قانون توزيع متَّجه عشوائي Distribution Law of a Random Vector )

ليكن $\mathbb{X}_n$ متَّجهاً عشوائياً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، فعندئذ يكون:

$$P_{\mathbb{X}_n}(B) = \left(P \circ \mathbb{X}_n^{-1}\right)(B) = P\left(\mathbb{X}_n^{-1}(B)\right) \qquad [5,29]$$

هو القياس المولَّد من $\mathbb{X}_n$ ، ويبرهن بسهولة على أنّه قياس احتمالي على $\left[\mathbb{R}^n, \mathfrak{R}^n\right]$. إنَّ القياس $P_{\mathbb{X}_n}$ يُدعى **قانون توزيع المتَّجه العشوائي** $\mathbb{X}_n$ .

**(٥,٤,٢,٢) ملاحظة**

بما أنّه من أجل أي متَّـجه عشوائي $\mathbb{X}_n$ فوق $[\Omega, \mathscr{A}, P]$ يوجد $n$ متغيِّر عشوائي $X_1$ ، $X_2$ ،... و $X_n$ بحيث يكون $\mathbb{X}_n = (X_1, X_2, ..., X_n)$ ، فإنَّ قانون توزيع $\mathbb{X}_n$ يُسمَّى في بعض الأحيان **قانون التوزيع المشترك** للمتغيِّرات العشوائيّة $X_1$ ، $X_2$ ،... و $X_n$ ويُرمز له عندئذ بـ $P_{X_1, X_2, ..., X_n}$ ، ولذلك من أجل $B := B_1 \times B_2 \times ... \times B_n$ من $\mathfrak{R}^n$ يمكننا أن نكتب ما يلي:

$$P_{\mathbb{X}_n}(B) = P_{X_1, X_2, ..., X_n}(B_1 \times B_2 \times ... \times B_n) = P(X_1 \in B_1, X_2 \in B_2, ..., X_n \in B_n) \qquad [5,30]$$

الآن، وبعد أن قمنا بتعريف قانون توزيع متَّـجه عشـوائي $\mathbb{X}_n$ يمكننا الانتقال إلى المرحلة التالية والمتمثِّلة بتعيين دالَّة توزيع المتَّجه العشوائي $\mathbb{X}_n$ .

**(٥,٤,٢,٣) تعريف (دالَّة توزيع متَّـجه عشوائي)**

ليكن $\mathbb{X}_n$ متَّجهاً عشوائياً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ ، ولنعرِّف من أجل هذا المتَّـجه العشوائي دالَّة حقيقية $F_{\mathbb{X}_n}$ على $\mathbb{R}^n$ كما يلي:

$$F_{\mathbb{X}_n} : \mathbb{R}^n \longrightarrow \mathbb{R} \; ; \vec{x}^n = (x_1, x_2, ..., x_n) \mapsto F_{\mathbb{X}_n}(\vec{x}^n) := P\left(\mathbb{X}_n \in \mathop{\times}_{k=1}^{n} (-\infty, x_k)\right) \qquad [5,31]$$

فعندئذ يُطلق على الدالَّة $F_{\mathbb{X}_n}$ اسم **دالَّة توزيع المتَّـجه العشوائي** $\mathbb{X}_n$ .

**(٥,٤,٢,٤) ملاحظات:**

١- في حال أخذنا $\mathbb{X}_n = (X_1, X_2, ..., X_n)$ فإنَّه يمكن كتابة العلاقة [5,31] على النحو الآتي أيضاً:

$$F_{\mathbb{X}_n}(\vec{x}^n) = P(X_1 < x_1, X_2 < x_2, ..., X_n < x_n) \qquad [5,32]$$

وذلك لأنَّ:

$$\begin{aligned} F_{\mathbb{X}_n}(\vec{x}^n) &= P\left(\mathbb{X}_n \in \mathop{\times}_{k=1}^{n} (-\infty, x_k)\right) = P\left(\mathbb{X}_n \in (-\infty, x_1) \times (-\infty, x_2) \times ... \times (-\infty, x_n)\right) \\ &= P\left((X_1, X_2, ..., X_n) \in (-\infty, x_1) \times (-\infty, x_2) \times ... \times (-\infty, x_n)\right) \\ &= P\left(X_1 \in (-\infty, x_1) \wedge X_2 \in (-\infty, x_2) \wedge ... \wedge X_n \in (-\infty, x_n)\right) \\ &= P(X_1 < x_1, X_2 < x_2, ..., X_n < x_n) \end{aligned}$$

٢- إنَّ دالَّة توزيع متَّـجه عشوائي $\mathbb{X}_n$ هي دالَّة التوزيع نفسها لقانون توزيـع $\mathbb{X}_n$ ؛ وذلك لأنَّه من أجل كل $\vec{x}^n \in \mathbb{R}^n$ لدينا:

$$\begin{aligned} F_{P_{\mathbb{X}_n}}(\vec{x}^n) &= P_{\mathbb{X}_n}((-\infty, x_1) \times (-\infty, x_2) \times ... \times (-\infty, x_n)) \\ &= P\left(\mathbb{X}_n^{-1}\left((-\infty, x_1) \times (-\infty, x_2) \times ... \times (-\infty, x_n)\right)\right) \\ &= P\left(\mathbb{X}_n \in (-\infty, x_1) \times (-\infty, x_2) \times ... \times (-\infty, x_n)\right) = F_{\mathbb{X}_n}(\vec{x}^n) \end{aligned}$$

٣- بما أنَّ $\mathbb{X}_n$ يمكن كتابته على الشكل $(X_1, X_2, ..., X_n)$ دوماً، فإنَّه يُطلق على الدالَّة $F_{\mathbb{X}_n}$ اسم **دالَّة التوزيع المشتركة**

Joint Distribution Function للمتغيِّرات العشوائيّة $X_1$ ، $X_2$ ، ... و $X_n$ ويُرمز لها حينئذ بـ $F_{X_1,X_2,...,X_n}$ ، ومن ثمَّ يُمكننا من أجل أي $\vec{x}^n = (x_1,x_2,...,x_n) \in \mathbb{R}^n$ أن نكتب الآتي:

$$F_{\mathbb{X}_n}(\vec{x}^n) = F_{X_1,X_2,...,X_n}(x_1,x_2,...,x_n) = P(X_1 < x_1, X_2 < x_2, ..., X_n < x_n) \quad [5,33]$$

٤- إنَّ تعريف دالَّة توزيع متَّجه عشوائي هو حالة خاصّة من تعريف دالَّة التوزيع الاحتمالية لقياس احتمالي (كما ذكر ذلك من أجل متغيِّر عشوائي في الفقرة / ٣/ من (٥,٢,٢,٢))، فلو كان $S$ قياساً احتمالياً على $[\mathbb{R}^n, \Re^n]$ ، وعرّفنا من أجله دالَّة حقيقية $F_S$ على $\mathbb{R}^n$ كما يلي:

$$F_S : \mathbb{R}^n \longrightarrow \mathbb{R} \; ; \; \vec{x}^n = (x_1,x_2,...,x_n) \mapsto F_S(\vec{x}^n) := S\left( \underset{k=1}{\overset{n}{\times}} (-\infty, x_k) \right) \quad [5,34]$$

فعندئذ (وبطريقة مماثلة لما سبق في الملاحظة / ٣/ من الفقرة (٥,٢,٢,٢)) تنتج دالَّة التوزيع الاحتمالية $F_{\mathbb{X}_n}$ من الدالَّة $F_S$ عندما نأخذ $S = P_{\mathbb{X}_n}$ حيث يمكن تبيان أنَّ $F_{P_{\mathbb{X}_n}} = F_{\mathbb{X}_n}$.

### (٥,٤,٣) خصائص دوال التوزيع الاحتمالية على $\mathbb{R}^n$

فيما يلي نقدِّم مبرهنة (وسنقبلها دون برهان، ولمن يود الاطلاع على البرهان يمكنه الرجوع إلى [64] ) تزودنا بأهمّ الخصائص المميِّزة لدوال التوزيع الاحتمالية على $\mathbb{R}^n$ .

#### (٥,٤,٣,١) مبرهنة

لتكن $F_{\mathbb{X}_n}$ دالَّة توزيع متَّجه عشوائي $\mathbb{X}_n$ فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ ، فعندئذ من أجل كل $\vec{x}^n = (x_1,x_2,...,x_n)$ من $\mathbb{R}^n$ يكون ما يلي مُحقَّقاً:

$\mathbf{e}_1^*$- الدالَّة $F_{\mathbb{X}_n}$ غير متناقصة على كل مركِّبة $x_k$ من مركبات المتَّجه $\vec{x}^n$ .

$\mathbf{e}_2^*$- إنَّ السلوك الحديّ لدالَّة التوزيع $F_{\mathbb{X}_n}$ يُحقِّق العلاقتين الآتيتين:

أ- إذا وجد على الأقل $k \in \mathbb{N}_n$ بحيث يُمكن للمُركَّبة $x_k$ أن تسعى إلى $-\infty$ (وسنرمز لذلك بـ $x_k \downarrow -\infty$ ) فعندئذ سيكون لدينا:

$$\lim_{x_k \downarrow -\infty} F_{\mathbb{X}_n}(x_1,...,x_{k-1},x_k,x_{k+1},...,x_n) = 0 \quad [5,35]$$

ب- إذا كان من أجل كل $k \in \mathbb{N}_n$ بحيث يمكن للمُركَّبة $x_k$ أن تسعى إلى $+\infty$ (وسنرمز لذلك بـ $x_k \uparrow +\infty$ ) فعندئذ سيكون لدينا:

$$\lim_{\substack{x_1 \uparrow +\infty \\ x_2 \uparrow +\infty \\ \vdots \\ x_n \uparrow +\infty}} F_{\mathbb{X}_n}(x_1,x_2,...,x_n) = 1 \quad [5,36]$$

$\mathbf{e}_3^*$- الدالَّة $F_{\mathbb{X}_n}$ مستمرَّة من اليسار عند كل مُركَّبة $x_k$ من مُركَّبات المتَّجه $\vec{x}^n$ .

#### (٥,٤,٣,٢) نتائج

١- من أجل كل $-\infty < a_1 < b_1 \leq +\infty$ و $-\infty < a_2 < b_2 \leq +\infty$ لدينا ما يلي مُحقَّقٌ:

$$F_{\mathbb{X}_n}(b_1,b_2) - F_{\mathbb{X}_n}(a_1,b_2) - F_{\mathbb{X}_n}(b_1,a_2) + F_{\mathbb{X}_n}(a_1,a_2) \geq 0 \quad [5,37\text{-}a]$$

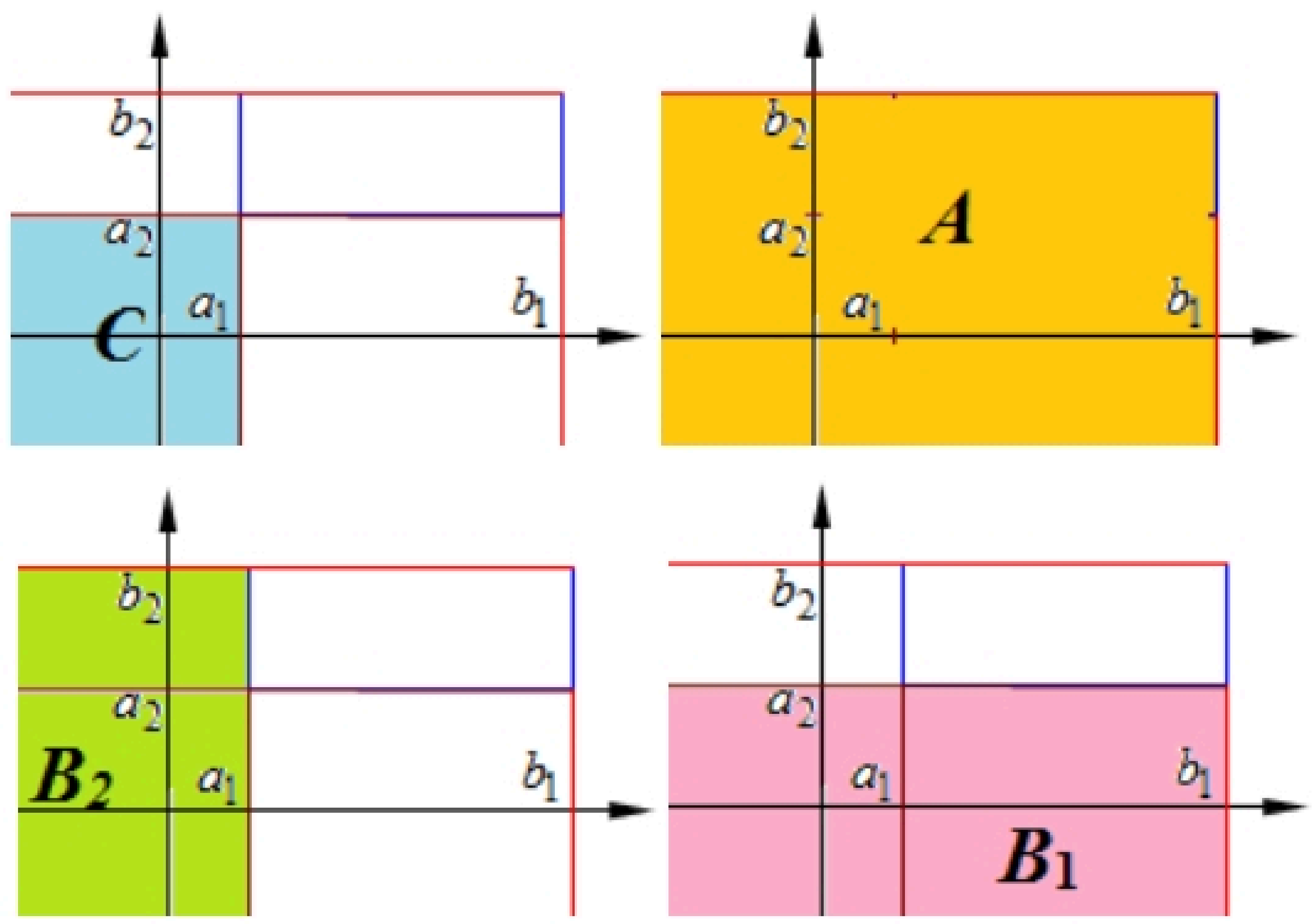
وهذه العلاقة تكافئ العلاقة الآتية من أجل كل $x_1, x_2 \in \mathbb{R}$ و $0 < h_2, h_1$ :

$$F_{\mathbb{X}_n}(x_1+h_1, x_2+h_2) - F_{\mathbb{X}_n}(x_1+h_1, x_2) - F_{\mathbb{X}_n}(x_1, x_2+h_2) + F_{\mathbb{X}_n}(x_1, x_2) \geq 0 \qquad [5,37\text{-}b]$$

**الإثبات**: لنقم بتعريف المجموعات (الساحات) الآتية:

$$A := (-\infty, b_1) \times (-\infty, b_2) \quad \& \quad C := (-\infty, a_1) \times (-\infty, a_2)$$
$$B_1 := (-\infty, b_1) \times (-\infty, a_2) \quad \& \quad B_2 := (-\infty, a_1) \times (-\infty, b_2)$$

والأشكال الأربعة الآتية (**الشكل** (٥,٢٤)) تعرض لنا هذه المجموعات.




الشكل (٥,٢٤)

فعندئذ لدينا ما يلي:

$$0 \leq P_{\mathbb{X}_n}\left([a_1, b_1) \times [a_2, b_2)\right) = P_{\mathbb{X}_n}\left(\{(-\infty, b_1) \setminus (-\infty, a_1)\} \times \{(-\infty, b_2) \setminus (-\infty, a_2)\}\right)$$
$$= P_{\mathbb{X}_n}\left((A \setminus (B_1 \cup B_2)) \cup C\right) = P_{\mathbb{X}_n}(A) - P_{\mathbb{X}_n}(B_1) - P_{\mathbb{X}_n}(B_2) + P_{\mathbb{X}_n}(C)$$

ومنه يمكننا أن نكتب الآتي:

$$P_{\mathbb{X}_n}\left([a_1, b_1) \times [a_2, b_2)\right) = P_{\mathbb{X}_n}\left((-\infty, b_1) \times (-\infty, b_2)\right) - P_{\mathbb{X}_n}\left((-\infty, b_1) \times (-\infty, a_2)\right)$$
$$- P_{\mathbb{X}_n}\left((-\infty, a_1) \times (-\infty, b_2)\right) + P_{\mathbb{X}_n}\left((-\infty, a_1) \times (-\infty, a_2)\right)$$
$$= F_{\mathbb{X}_n}(b_1, b_2) - F_{\mathbb{X}_n}(b_1, a_2) - F_{\mathbb{X}_n}(a_1, b_2) + F_{\mathbb{X}_n}(a_1, a_2)$$

وبهذا يتمُّ الإثبات.

٢- من النتيجة السّابقة يمكننا أن نكتب المتباينة الآتية من أجل متَّجه عشوائي $\mathbb{X}_2 = (X_1, X_2)$ :

$$P\left(a_1 \leq X_1 < b_1, a_2 \leq X_2 < b_2\right) = F_{X_1, X_2}(b_1, b_2) - F_{X_1, X_2}(a_1, b_2) - F_{X_1, X_2}(b_1, a_2) + F_{X_1, X_2}(a_1, a_2) \geq 0 \qquad [5,38]$$

علماً أنّ $-\infty < a_1 < b_1 \leq +\infty$ و $-\infty < a_2 < b_2 \leq +\infty$ .

سنعرض فيما يلي الشروط الكافية لكي تكون دالّة حقيقية $F$ على $\mathbb{R}^n$ دالّة توزيع احتمالية، وسنمهد لذلك بالتعريف الآتي.

**(٥,٤,٣,٣) تعريف**

لنرمز بـ $\mathbf{F}(\mathbb{R}^n)$ لمجموعة كل الدوال الحقيقية المعرَّفة على $\mathbb{R}^n$، ولنأخذ $0<h$ عدد حقيقي، ولنعرِّف من أجل كل $k \ni \mathbb{N}_n$ مؤثِّراً Operator $\Delta_h^k$ على $\mathbf{F}(\mathbb{R}^n)$ كما يلي:

$$\begin{aligned} \Delta_h^k : \mathbf{F}(\mathbb{R}^n) &\longrightarrow \mathbf{F}(\mathbb{R}^n) \\ f \mapsto \Delta_h^k(f) &:= f(x_1,\ldots,x_{k-1},x_k+h,x_{k+1},\ldots,x_n) - f(x_1,\ldots,x_k,\ldots,x_n) \end{aligned} \quad [5,39]$$

بمساعدة هذا التعريف يمكننا الحصول على النتيجة المهمّة الآتية.

**(٥,٤,٣,٤) نتيجة**

لتكن $F_{\mathbb{X}_2}$ دالّة توزيع متّجه عشوائي $\mathbb{X}_2$، فعندئذ من أجل كل $x_1, x_2 \ni \mathbb{R}$ و $0<h_2,h_1$ ستكون المتباينة الآتية مُحقَّقة:

$$\Delta_{h_2}^2 \Delta_{h_1}^1 F_{\mathbb{X}_2}(x_1,x_2) \geq 0 \quad [5,40]$$

وذلك لأنّ:

$$\begin{aligned} \Delta_{h_2}^2\left(\Delta_{h_1}^1 F_{\mathbb{X}_2}(x_1,x_2)\right) &= \Delta_{h_2}^2\left(F_{\mathbb{X}_2}(x_1+h_1,x_2) - F_{\mathbb{X}_2}(x_1,x_2)\right) \\ &= \Delta_{h_2}^2\left(F_{\mathbb{X}_2}(x_1+h_1,x_2)\right) - \Delta_{h_2}^2\left(F_{\mathbb{X}_2}(x_1,x_2)\right) \\ &= F_{\mathbb{X}_2}(x_1+h_1,x_2+h_2) - F_{\mathbb{X}_2}(x_1+h_1,x_2) - F_{\mathbb{X}_2}(x_1,x_2+h_2) - F_{\mathbb{X}_2}(x_1,x_2) \geq 0 \end{aligned}$$

سنقبل المبرهنة الآتية دون برهان حيث تُعدّ تعميماً للنتيجة السّابقة.

**(٥,٤,٣,٥) مبرهنة**

لتكن $F_{\mathbb{X}_n}$ دالّة توزيع متّجه عشوائي $\mathbb{X}_n$ مع $n \ni \mathbb{N}$ مثبّت، فعندئذ دالّة التوزيع $F_{\mathbb{X}_n}$ تُحقّق الخاصيّة الآتية:

$\mathbf{e}_4^*$ – من أجل أيّ متّجه $(x_1,x_2,\ldots,x_n) \ni \mathbb{R}^n$ وأيّة أعداد حقيقية موجبة $h_1$ ، $h_2$ ، … و $0<h_n$ ، سيكون لدينا:

$$\Delta_{h_n}^n \Delta_{h_{n-1}}^{n-1} \ldots \Delta_{h_2}^2 \Delta_{h_1}^1 F_{\mathbb{X}_2}(x_1,x_1,\ldots,x_n) \geq 0 \quad [5,41]$$

**(٥,٤,٣,٦) ملاحظات**

١ – إنَّ المتباينة السّابقة [5,38] يمكن تعميمها على $\mathbb{R}^n$ إلاّ أنَّ الشكل النّاتج سيكون أكثر تعقيداً، علماً أنّه يمكن استنتاجها من خلال الخاصية $\mathbf{e}_4^*$ السّابقة.

٢ – إنَّ أيّة دالّة توزيع احتمالية على $\mathbb{R}^n$ تُحقّق الخواص $\mathbf{e}_1^*$ و $\mathbf{e}_2^*$ و $\mathbf{e}_3^*$، ولكن العكس ليس صحيحاً، أي إنّه ليست كل دالّة حقيقية معرّفة على $\mathbb{R}^n$ وتُحقّق الخواص $\mathbf{e}_1^*$ و $\mathbf{e}_2^*$ و $\mathbf{e}_3^*$ هي دالّة توزيع احتمالية، فلو أخذنا على سبيل المثال الدالّة الحقيقية $F$ والمعرَّفة على $\mathbb{R}^2$ كما يلي:

$$F(x,y) := \begin{cases} 1 & for \;\; x>0,\, y>1 \;\; or \;\; x>2,\, y>0 \\ 0 & otherwise. \end{cases}$$

فنجد أنَّ هذه الدالَّة تُحقِّق الخواص $\mathbf{e}_1^*$ و $\mathbf{e}_2^*$ و $\mathbf{e}_3^*$، ولكن نلاحظ أنَّه من أجل $h_1 = h_2 = 2$ و $x_1 = x_2 = 1$ لدينا الآتي:

$$F(3,3)-F(3,1)-F(1,3)+F(1,1)=1-1-1+0=-1<0$$

وهذا يعني أنَّ الدالَّة $F$ لا تُحقِّق المتباينة [5,37-a] من أجل النقطة (1,1) والقيمتين $h_1 = h_2 = 2$، ومن ثمَّ فإنَّها لا تصلح أن تكون دالَّة توزيع احتمالية على $\mathbb{R}^2$.

٣- لقد بيَّنت لنا الملاحظة السّابقة أنَّ خصائص دوال التوزيع للمتَّجهات العشوائيّة قد تختلف عن خصائص دوال التوزيع للمتغيِّرات العشوائيّة الوحيدة البُعد، ففي الواقع توجد عدَّة فوارق بين دوال التوزيع للمتَّجهات العشوائيّة ودوال التوزيع للمتغيِّرات العشوائيّة ذات البُعد الواحد سنذكر منها فارقين مهمَّين هما:

أ- نلاحظ أنَّ الخاصيَّة $\mathbf{e}_1^*$ تنتج من الخاصيَّة $\mathbf{e}_4^*$، ومن ثمَّ فإنَّ العلاقات $\mathbf{e}_2^*$ و $\mathbf{e}_3^*$ و $\mathbf{e}_4^*$، تُميِّز دالَّة التوزيع الاحتمالية بشكل تام، بمعنى أنَّ كل دالَّة توزيع احتمالية $F_{\mathbb{X}_n}$ تُحقِّق الخواص $\mathbf{e}_2^*$ و $\mathbf{e}_3^*$ و $\mathbf{e}_4^*$، والعكس صحيح أيضاً، أي إنَّه إذ كانت $F$ دالَّة حقيقية معرَّفة على $\mathbb{R}^n$ وتُحقِّق الخواص $\mathbf{e}_2^*$ و $\mathbf{e}_3^*$ و $\mathbf{e}_4^*$ فإنَّ هذه الدالَّة $F$ ستكون دالَّة توزيع احتمالية على $\mathbb{R}^n$، بمعنى أنَّه يوجد متَّجه عشوائي $\mathbb{X}_n$ بحيث تكون $F$ هي دالَّة توزيـعه الاحتمالية.

ب- نعلم أنَّه إذا كان $X$ متغيِّراً عشوائياً مستمرّاً فإنَّ دالَّة توزيعه $F_X$ تكون مستمرّة على $\mathbb{R}$، ولكنَّ هذه الخاصية تصبح غير صحيحة من أجل المتَّجهات العشوائيّة، وذلك لأنَّه توجد دوال توزيع احتمالية لمتَّجهات عشوائيَّة مستمرّة (انظر تعريف المتَّجه العشوائي المستمرّ في فقرة تصنيف المتَّجهات العشوائيّة) تعاني من نقط انقطاع، فعلى سبيل المثال لو أخذنا $\mathbb{X}_2 = (X_1, X_2)$ متَّجهاً عشوائياً بدالَّة توزيع $F_{\mathbb{X}_2}$ مُعطاة من خلال العلاقة الآتية:

$$F_{\mathbb{X}_2}(x_1,x_2)=\begin{cases} 0 & for\ x_1 \le 0 \quad or \quad x_2 \le 0 \\ x_1 \cdot x_2 & for\ 0 < x_1 \le 1 \ and \ 0 < x_2 \le \frac{1}{2} \\ \frac{1}{2}x_1 & for\ 0 < x_1 \le 1 \ and \ x_2 > \frac{1}{2} \\ 2x_2 & for\ x_1 > 1 \quad and \ 0 < x_2 \le \frac{1}{2} \\ 1 & for\ x_1 > 1 \quad and \ x_2 > \frac{1}{2} \end{cases}$$

فنجد أنَّ للمتَّجه العشوائي $\mathbb{X}_2$ احتمال يساوي $\frac{1}{2}$ عند النقاط $x_1 = 1$ و $0 \le x_2 \le \frac{1}{2}$ وذلك لأنَّ:

$$P\left(X_1 = 1\ ,\ 0 \le X_2 \le \frac{1}{2}\right) = F_{\mathbb{X}_2}\left(1\,,\frac{1}{2}\right) - F_{\mathbb{X}_2}(1\,,0) = \frac{1}{2} - 0 = \frac{1}{2}$$

ومن جهة أخرى واضح أنَّ كل نقطة ذات إحداثي $(1, x_2)$ مع $\frac{1}{2} < x_2 < \infty$ هي نقطة انقطاع لدالَّة التوزيع $F_{\mathbb{X}_2}$ رغم أنَّ المتَّجه العشوائي $\mathbb{X}_2$ ليس له قفزات على الفترة $\frac{1}{2} < x_2 < \infty$ وعند $x_1 = 1$.

### (٥ , ٤ , ٤) تصنيف المتَّجهات العشوائيَّة Classification of Random Vectors

لقد قمنا فيما سبق بدراسة موجزة نسبياً لدوال التوزيع الاحتمالية لمتَّجهات عشوائية، ومن ثَّم أوضحنا بعض الخصائص المهمّة لهذه الدالّة، ولكن لدى دراسة الكثير من الظواهر والتجارب العشوائيّة تصادفنا متَّجهات عشوائيَّة مختلفة الأصناف فمنها ذات مركبات لمتغيِّرات عشوائيَّة متقطِّعة، وأخرى ذات مركبات لمتغيِّرات عشوائيَّة مستمرّة، وثالثة ذات مركبات لمتغيِّرات عشوائيَّة متقطِّعة ومستمرّة أو غير مستمرّة. لذلك نجد أنَّه لا بدَّ من التمييز بين هذه المتَّجهات العشوائيّة لتبسيط دراستها وتوضيح العمليّات الحسابية بشأنَّها، وفي كتابنا هذا سنركِّز دراستنا على نموذجين فقط وهما:

- المتَّجهات العشوائيّة المتقطِّعة،
- المتَّجهات العشوائيّة المستمرّة.

#### (٥ , ٤ , ٤ , ١) المتَّجهات العشوائيَّة المتقطِّعة Discrete Random Vectors

على غرار ما قمنا به من أجل المتغيِّرات العشوائيّة الوحيدة البعد سنقوم أولاً بتعريف المتَّجه العشوائيّ المتقطِّع الذي يقدِّمه لنا التعريف الآتي.

#### (٥ , ٤ , ٤ , ١ , ١) تعريف (المتَّجه العشوائي المتقطِّع)

ليكن $\mathbb{X}_n$ متَّجهاً عشوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فإذا كانت جميع مركبات هذا المتَّجه هي متغيِّرات عشوائيَّة متقطِّعة، فعندئذ يُقال عن $\mathbb{X}_n$ إنَّه **متَّجه عشوائي متقطِّع**.

#### (٥ , ٤ , ٤ , ١ , ٢) ملاحظات

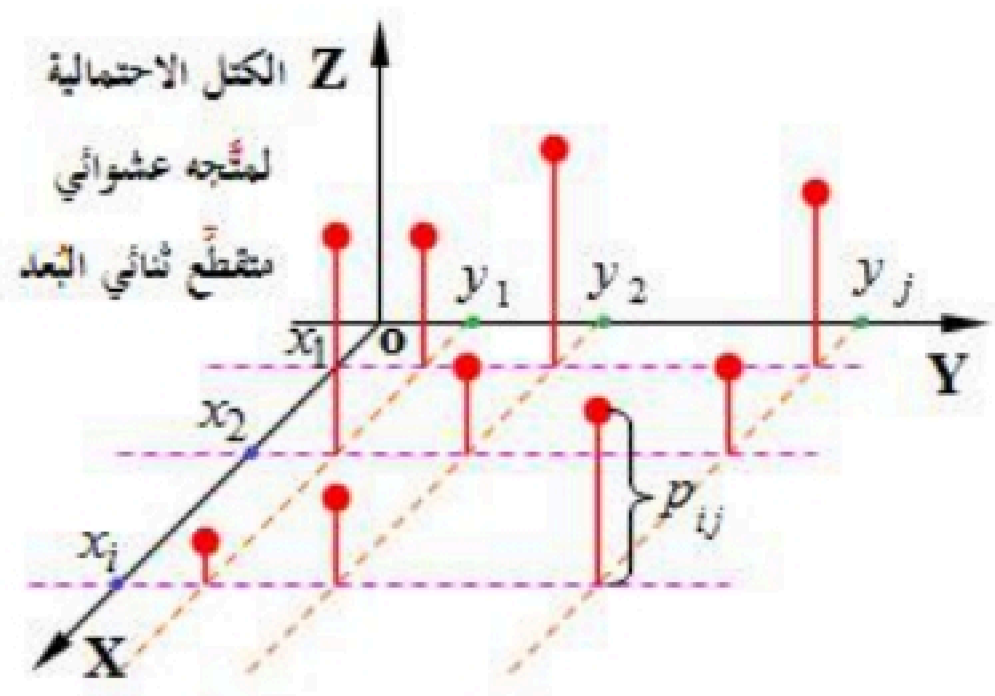
١- بفرض أنَّ $\mathbb{X}_n = (X_1, X_2, \dots, X_n)$ متَّجه عشوائي متقطِّع بحيث إنَّه من اجل كل $k \in \mathbb{N}_n$ للمركَّبة $X_k$ مجموعة القيم $\mathbf{X}_k = \{x_{k\,i_k} \; ; \; i_k \in I_k \subseteq \mathbb{N}\}$ ، فعندئذ تُسمّى النقاط $(x_{1i_1}, x_{2i_2}, \dots, x_{n\,i_n})$ **مواضع قفزات للمتَّجه العشوائي** $\mathbb{X}_n$، في حين أنَّ قيمة احتمال أنَّ يأخذ المتَّجه العشوائي $\mathbb{X}_n$ النقطة $(x_{1i_1}, x_{2i_2}, \dots, x_{n\,i_n})$ تُدعى **الكتلة الاحتمالية للمتَّجه العشوائي** $\mathbb{X}_n$ عند هذا النقطة، ويُرمز لهذه الكتلة الاحتمالية بـ $p_{i_1 i_2 \dots i_n}$، وهذه القيمة تساوي ارتفاع القفزة للمتَّجه العشوائي $\mathbb{X}_n$ عند الموضع المذكور آنفاً، أي إنَّه لدينا:

$$p_{i_1 i_2 \dots i_n} := P\left(X_1 = x_{1i_1}, X_2 = x_{2i_2}, \dots, X_n = x_{n\,i_n}\right) \qquad [5,42]$$

وهنا يجب أن تكون العلاقة الآتية مُحقَّقة:

$$\sum_{i_1 \in I_1 ;\, i_2 \in I_2 ;\, \dots ;\, i_n \in I_n} p_{i_1 i_2 \dots i_n} = 1 \qquad [5,43]$$

٢- عندما يكون بُعد المتَّجه $n \geq 3$ فإنَّه على سبيل التبسيط يُفضّل استخدام الرموز $i$ و$j$ و$k$ عوضاً $1i_1$ و $2i_2$ و $3i_3$ على الترتيب، وفي هذه الحالة الخاصّة يُفضّل استخدام الأحرف $X$ و$Y$ و$Z$ عوضاً عن $X_1$ و $X_2$ و $X_3$ على الترتيب، فعلى سبيل المثال يعرض لنا الشكل (٥ , ٢٥) الآتي تمثيلاً بيانياً توضِّح قيم الاحتمالات الكتلية لمتَّجه عشوائي متقطِّع $\mathbb{X}_2 = (X, Y)$. حيث افترضنا أنَّ $\mathbf{X} = \{x_i \; ; \; i \in I \subseteq \mathbb{N}\}$ و $\mathbf{Y} = \{y_j \; ; \; j \in J \subseteq \mathbb{N}\}$ هي مجموعة قيم المركَّبة $X$ و$Y$ على الترتيب، فنلاحظ أنَّ الكتل الاحتمالية للمتَّجهات العشوائيّة المتقطِّعة كائنة في نقاط من الفضاء الثلاثي الأبعاد (أي على شكل نقطة بوزن يساوي كتلة الاحتمال) وترتفع عن المستوي **XoY** وفقاً لقيم لهذه الكتل الاحتمالية.




الشكل (٥, ٢٥)

٢- من تعريف دالَّة التوزيع لمتَّجه عشوائي يُمكننا من أجل كل $(x_1,x_2,...,x_n) \in \mathbb{R}^n$ أن نصيغ دالَّة توزيع متَّجه عشوائي متقطِّع $\mathbb{X}_n$ من خلال العرض الآتي:

$$\begin{aligned} F_{\mathbb{X}_n}(x_1,x_2,...,x_n) &= P\left(X_1<x_1,X_2<x_2,...,X_n<x_n\right) \\ &= \sum_{i_1\in I_1;x_{1i_1}<x_1}\ \sum_{i_2\in I_2;x_{2i_2}<x_2}\cdots\sum_{i_n\in I_n;x_{ni_n}<x_n} P\left(X_1=x_{1i_1},X_2=x_{2i_2},...,X_n=x_{ni_n}\right) \\ &= \sum_{i_1\in I_1;x_{1i_1}<x_1}\ \sum_{i_2\in I_2;x_{2i_2}<x_2}\cdots\sum_{i_n\in I_n;x_{ni_n}<x_n} p_{i_1i_2....i_n} \end{aligned}$$

### (٥, ٤, ٤, ٢) المتَّجهات العشوائيّة المستمرّة Continuous Random Vectors

لقد لاحظنا سا بقاً أنَّ المتغيِّرات العشوائيّة المستمرّة وحيدة البعد تكون مميَّزة من خلال دوال كثافاتها الاحتمالية، وكذلك الأمر بالنسبة إلى المتَّجهات العشوائيّة المستمرّة، والتعريف الآتي يبيّن لنا ذلك.

#### (٥, ٤, ٤, ٢, ١) تعريف (المتَّجه العشوائي المستمرّ)

ليكن $\mathbb{X}_n$ متَّجهاً عشوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فإذا وجدت دالَّة حقيقية $f$ غير سالبة معرَّفة وكمولة على $\mathbb{R}^n$، وبحيث إنه من أجل كل $(x_1,x_2,...,x_n) \in \mathbb{R}^n$ يكون لدينا ما يلي مُحَقَّقاً:

$$F_{\mathbb{X}_n}(x_1,x_2,...,x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} f(t_1,t_2,...,t_n)\ dt_1dt_2...dt_n$$

فعندئذ يُقال إنَّ $\mathbb{X}_n$ هو متَّجه عشـوائي مستمرّ بدالَّة كثافة احتمالية $f$.

إنَّ الدالَّة $f$ والمحقِّقة للشروط المذكورة أعلاه هي في الواقع كثافة رادون- نيكوديم من $P_{\mathbb{X}_n}$ (بالنسبة إلى قياس لوبيغ $\lambda$)، ولذلك يُرمز لها بـ $f_{\mathbb{X}_n}$، أي إنَّه لدينا:

$$F_{\mathbb{X}_n}(x_1,x_2,...,x_n) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2}\cdots\int_{-\infty}^{x_n} f_{\mathbb{X}_n}(t_1,t_2,...,t_n)\ dt_1dt_2...dt_n \qquad [5,45]$$

ويكون لدينا في هذه الحالة (تنتج مباشرة من خواص دالَّة التوزيع الاحتمالية $F_{\mathbb{X}_n}$ ):

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{\mathbb{X}_n}(x_1,x_2,\ldots,x_n)\, dx_1 dx_2 \ldots dx_n = 1$$

### (٥,٤,٤,٢,٢) ملاحظات

١ - إنَّ الدالَّة $f_{\mathbb{X}_n}$ تدعى في بعض الأحيان **دالَّة الكثافة المشتركة** Joint Density Function لـ $X_1$ ، $X_2$ ،... و $X_n$ ، ولذلك تُكتب بالشكل الآتي أيضاً:

$$f_{\mathbb{X}_n}(\vec{x}^{\,n}) = f_{X_1,X_2,\ldots,X_n}(x_1,x_2,\ldots,x_n) \quad ; \vec{x}^{\,n} = (x_1,x_2,\ldots,x_n)$$

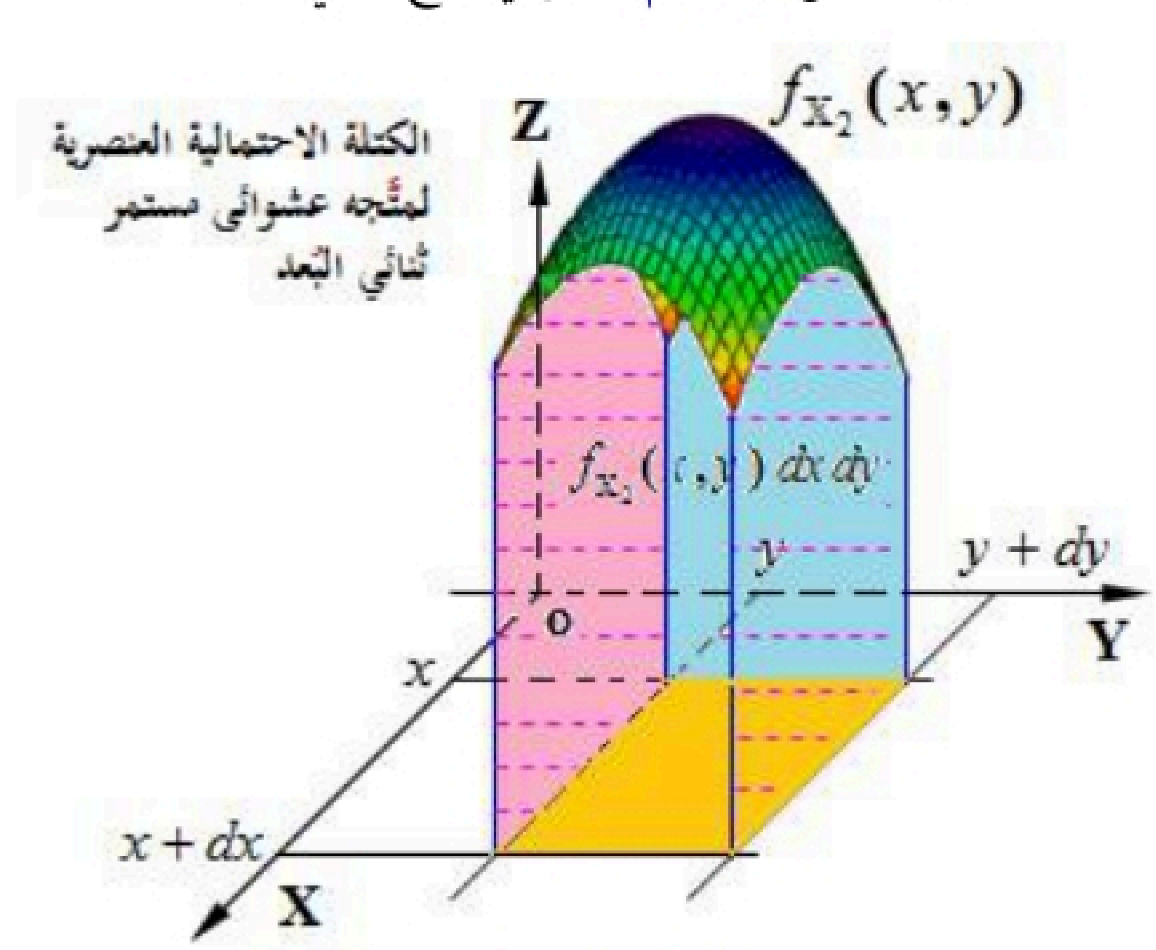
٢- إذا كانت الدالَّة $f_{\mathbb{X}_n}$ مستمرّة عند نقطة $\vec{x}^{\,n} = (x_1,x_2,\ldots,x_n)$ ∋ $\mathbb{R}^n$ ، فعندئذ ستكون دالَّة التوزيع $F_{\mathbb{X}_n}$ قابلة للاشتقاق الجزئي بالنسبة إلى كل مُركّبة $x_k$ من مُركّبات المتَّجه $\vec{x}^{\,n}$ ، ولدينا العلاقة الآتية مُحقَّقة:

$$\frac{\partial^n}{\partial x_1\, \partial x_2 \ldots \partial x_n} F_{\mathbb{X}_n}(x_1,x_2,\ldots,x_n) = f_{\mathbb{X}_n}(x_1,x_2,\ldots,x_n) \qquad [5,46]$$

وهذه العلاقة الأخيرة تبيِّن أنَّه من أجل أيّة قيم حقيقية موجبة $h_1$ ، $h_2$ ... و $0 < h_n$ ستكون العلاقة الآتية مُحقَّقة أيضاً:

$$f_{\mathbb{X}_2}(x_1,x_1,\ldots,x_n) = \lim_{\substack{h_1 \to 0 \\ h_2 \to 0 \\ \vdots \\ h_n \to 0}} \frac{P(x_1 \le X_1 < x_1 + h_1, x_2 \le X_2 < x_2 + h_2, \ldots, x_n \le X_n < x_n + h_n)}{h_1 \cdot h_2 \cdot \ldots \cdot h_n}$$

من أجل الحالة الخاصَّة $n = 2$ يوضِّح لنا الشكل الآتي (٥,٢٦) الرسم البياني الذي يُحدِّد قيمة الاحتمال العنصري $f_{\mathbb{X}_2}(x,y)\,dx\,dy$ لمتَّجه عشوائي مستمرّ $\mathbb{X}_2 = (X, Y)$ بدالَّة كثافة احتمالية $f_{\mathbb{X}_2}$ ، وقد أُعطي $x$ تزايداً قدره $dx$ في حين أُعطي $y$ تزايداً قدره $dy$ . حيث نلاحظ أنَّ $f_{\mathbb{X}_2}(x,y)\,dx\,dy$ التي تمثِّل الكثافة الاحتمالية للمتَّجه العشوائي المستمر $\mathbb{X}_2 = (X, Y)$ في المجال العنصري $[x, x+dx) \times [y, y+dy)$ لها شكل (أو حجم) كائن في فراغ ثلاثي الأبعاد.




الشكل (٥,٢٦)

### (٥,٤,٣) المتَّجهات العشوائيّة المختلطة Mixed Random Vectors

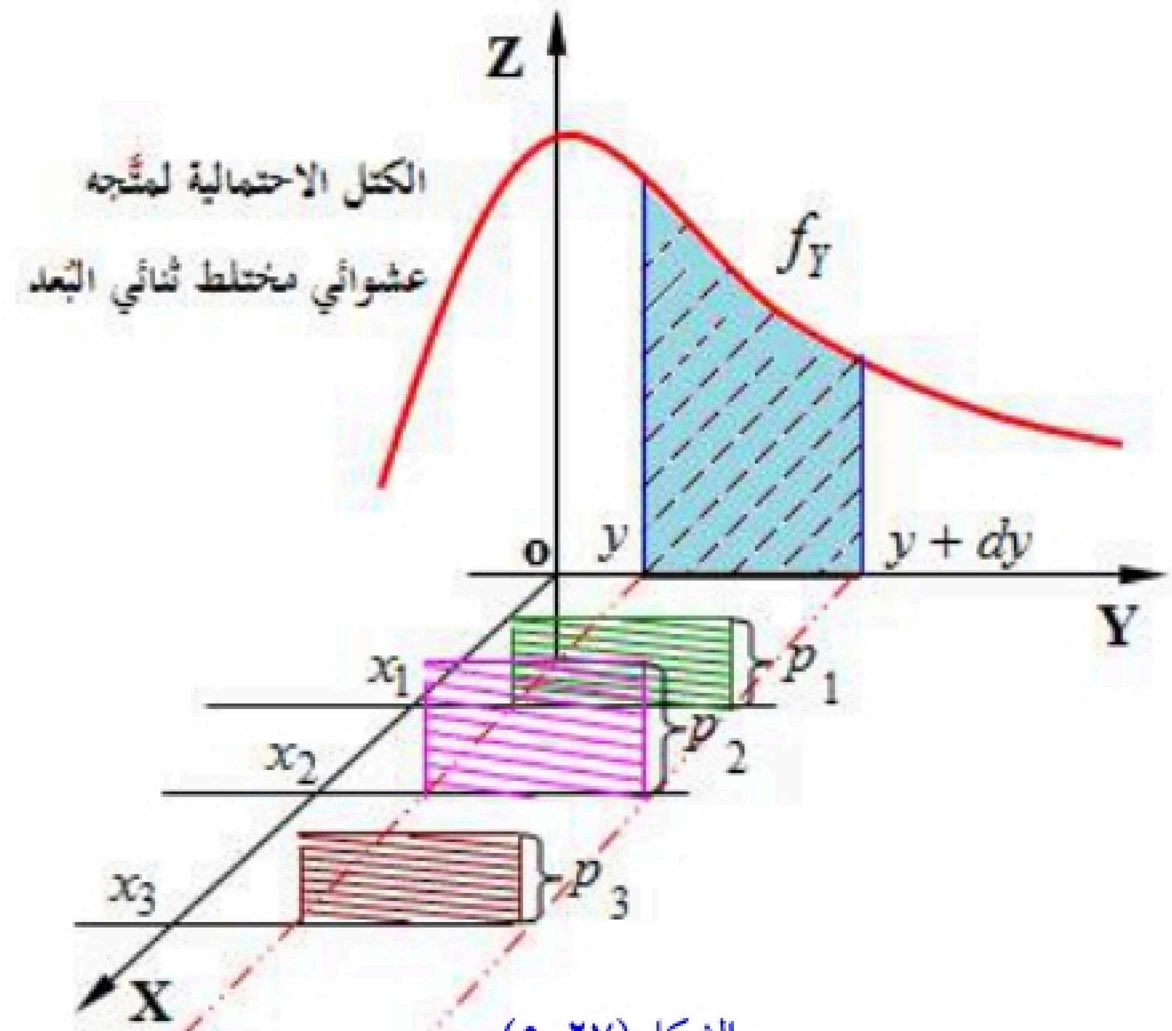
في الحقيقة تُعدّ المتَّجهات العشوائيّة المختلطة من أكثر المتَّجهات العشوائيّة ظهوراً وتكراراً في الدراسات التطبيقية والنظرية على

حدٍّ سواء، ولكنَّ دراستها في الحالة العامّة ليس سهلاً وعلاقاتها غير بسيطة، ولهذا سنكتفي بتعريفها وتقديم بعض الملاحظات حولها.

**(١, ٣, ٤, ٤, ٥) تعريف (المتَّجه العشوائي المختلط)**

ليكن $\mathbb{X}_2$ متَّجهاً عشوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فإذا كانت مركَّبات هذا المتَّجه العشوائي مكوَّنة من صنفين مختلفين من المتغيِّرات العشوائيّة (متقطِّعة – مستمرَّة – غير مستمرَّة) على الأقل، فعندئذ يُقال إنَّ $\mathbb{X}_2$ هو **متَّجه عشوائي مختلط**.

من أجل توضيح التوزيع الاحتمالي للمتَّجه العشوائي المختلط سنأخذ الحالة الخاصة عندما يكون لدينا $\mathbb{X}_2 = (X , Y)$ مع $X$ متغيِّر عشوائي متقطِّع بمجموعة قيم $\mathbf{X} := \{x_i \;;\; i \in I \subseteq \mathbb{N}\}$، و$Y$ متغيِّر عشوائي مستمرّ بدالَّة كثافة احتمالية $f_Y$ (انظر الشكل الجانبي (٥, ٢٧))، فعندئذ تتوزَّع الكثافة الاحتمالية على جزء عنصري بطول $dy$، وعندما يأخذ $X$ القيم $x_i$ فإنَّ الكثافة الاحتمالية المذكورة آنفاً هي تلك المساحة المستطيلة والقائمة على المستقيمات المارة من النقاط $x_i$ والموازية للمحور $\mathbf{oY}$.



الشكل (٥, ٢٧)

هنا يجب أن يكون مجموع مساحات هذه المستطيلات مساوياً للمساحة المظلَّلة والمحصورة من الأعلى بمنحنى الدالَّة $f_Y$، وهذه الكثافة الاحتمالية في الفترة العنصرية $[y , y+dy)$ والواقعة على المستقيم المار من $x_i$ والموازي للمحور $\mathbf{oY}$ تساوي المقدار الآتي:

$$f_{\mathbb{X}_2}(x_i , y)\, dy = P(X = x_i \;,\; y \le Y < y+dy) \qquad [5,47]$$

وهكذا نلاحظ أنَّه من أجل متَّجهات عشوائية مختلطة $\mathbb{X}_2 = (X , Y)$ مع $X$ متغيِّر عشوائي متقطِّع بمجموعة قيم $\mathbf{X} := \{x_i \;;\; i \in I \subseteq \mathbb{N}\}$، و$Y$ متغيِّر عشوائي مستمرّ بدالَّة كثافة احتمالية $f_Y$، فإنَّ كتلها الاحتمالية كائنة في مستويات موازية للمستوي $\mathbf{YoZ}$ في الفراغ الثلاثي الأبعاد (أي على شكل مساحات).

أمّا بشأن دالَّة التوزيع الاحتمالية $F_{\mathbb{X}_2}$ فإنَّها تكون في هذه الحالة غير مستمرَّة عند النقاط $x_i$ من أجل كل $I \ni i$، ولكنَّها مستمرَّة على المركَّبة $y$ فقط. ويكون لدينا:

$$F_{\mathbb{X}_2}(x,y) = \int_{-\infty}^{y} \sum_{i\,;\,x_i<x} f_{\mathbb{X}_2}(x_i , t)\; dt \qquad [5,48]$$

لاحظ أنَّه يمكننا المبادلة بين التكامل والمجموع في العلاقة السّابقة وذلك من أجل كل $\mathbb{R}^2 \ni (x , y)$.

بشكل مُماثل تُناقش الحالة التي يكون فيها لدينا $\mathbb{X}_2 = (X , Y)$ مع $X$ متغيِّر عشوائي مستمرّ بدالَّة كثافة احتمالية $f_X$، وأمّا $Y$ فهو متغيِّر عشوائي متقطِّع بمجموعة قيم $\mathbf{Y} := \{y_j \;;\; j \in J \subseteq \mathbb{N}\}$، حيث قيمة الكثافة الاحتمالية في الفترة العنصرية $[x , x+dx)$ والقائمة على المستقيم المار من $y_j$ والموازي للمحور $\mathbf{oX}$ ومحدودة من الأعلى بمنحني الدالَّة $f_X$ سوف يساوي:

$$f_{\mathbb{X}_2}(x \;,\; y_j)\, dx := P\left(x \le X < x+dx \;\;,\; Y = y_j\right) \qquad [5,49]$$

وتكون دالَّة التوزيع $F_{\mathbb{X}_2}$ غير مستمرَّة عند النقطة $y_j$ من أجل كل $J \ni j$، ولكنَّها مستمرَّة على المركَّبة $x$ فقط، ويكون لدينا:

$$F_{\mathbb{X}_2}(x,y)=\int_{-\infty}^{x}\sum_{j\,;\,y_j<y} f_{\mathbb{X}_2}(t,y_j)\,dt \qquad [5,50]$$

وهنا يمكن المبادلة بين التكامل والمجموع من أجل كل $(x\,,\,y)\in\mathbb{R}^2$ أيضاً.

**(٤ , ٤ , ٥) أمثلة**

١- ليكن $\mathbb{X}_2=(X\,,Y\,)$ متَّجهاً عشوائياً متقطِّعاً فوق الفضاء الاحتمالي لتجربة إلقاء حجر نرد متوازن لمرَّة واحدة فقط، ولنأخذ $X$ و $Y$ متغيِّرين عشوائيين فوق الفضاء الاحتمالي لهذه التجربة العشوائيّة ومعرَّفين من أجل كل $\omega\in\Omega$ كما يلي:

$$X(\omega)=\omega \qquad \& \qquad Y(\omega)=2\omega$$

ولنقم بتعيين دالَّة التوزيع الاحتمالية لهذا المتَّجه العشوائي، ومن ثَّم حسب قيمتها عند النقاط $(1,1)$ و $(3,7)$ و $(8,15)$.

الحل: من معطيات المثال لدينا مركَّبا الفضاء الاحتمالي $[\,\Omega,\mathscr{A},P\,]$ للتجربة العشوائيّة المُعطاة هي $\Omega=\mathbb{N}_6$ ولنأخذ $\mathscr{A}=2^{\Omega}$ فنجد أنَّ $P(\{\omega\})=\frac{1}{6}$ من أجل كل $\omega\in\Omega$، وكذلك لمجموعة قيم المتَّجه العشوائي $\mathbb{X}_2$ العرض الآتي:

$$\mathbf{G}:=\{(1,2),(2,4),(3,6),(4,8),(5,10),(6,12)\}$$

وبما أنَّ حجر النرد متوازن فإنَّ احتمالات الحوادث الابتدائية ستكون متساوية، ومن ثَّم ينتج لدينا ما يلي:

$$p_{ij}=P(X=x_i\,,Y=y_j)=\begin{cases}\frac{1}{6} & for \quad j=2i\\ 0 & otherwise\end{cases}$$

وأمَّا دالَّة توزيع هذا المتَّجه العشوائي فلها العرض الآتي:

$$F_{\mathbb{X}_2}(x,y)=\sum_{i\,;\,x_i<x}\ \sum_{j\,;\,y_j<y} P(X=x_i\,,Y=y_j)$$

ومنه يكون لدينا على سبيل المثال:

$$F_{\mathbb{X}_2}(1\,,1)=\sum_{i\,;\,i<1}\ \sum_{j\,;\,j<1} p_{ij}=0$$

$$F_{\mathbb{X}_2}(3\,,7)=\sum_{i\,;\,i<3}\ \sum_{j\,;\,j<7} p_{ij}=p_{12}+p_{24}=\frac{1}{6}+\frac{1}{6}=\frac{1}{3}$$

$$F_{\mathbb{X}_2}(8\,,15)=\sum_{i\,;\,i<8}\ \sum_{j\,;\,j<15} p_{ij}=p_{12}+p_{24}+p_{36}+p_{48}+p_{5\,10}+p_{6\,12}=1$$

علماً أنَّ الاحتمالات $p_{ij}$ ذات القيمة صفر تمَّ تجاوزها وعدم كتابتها في الصيغ السّابقة.

٢- ليكن $\mathbb{X}_2=(X\,,Y\,)$ متَّجهاً عشوائياً مستمرّاً بدالَّة كثافة احتمالية مُعطاة من خلال العلاقة:

$$f_{\mathbb{X}_2}(x,y)=\begin{cases}\alpha x\,y & for \ (x\,,y)\in A\\ 0 & otherwise\end{cases}$$

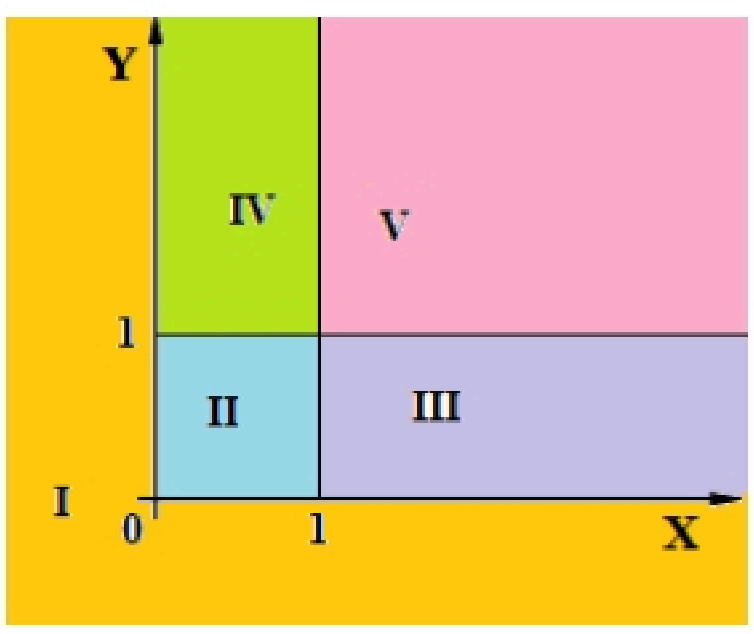
علماً أنَّ $A$ هو المجال $(0\,,1)\times(0\,,1)$ ، وأمّا $\alpha$ فهو ثابت عددي مغاير للصفر، ولنقم بتعيين قيمة الثابت $\alpha$ ، ومن ثمَّ تعيين دالَّة توزيع المتَّجه $\mathbb{X}_2=(X\,,Y\,)$ .

الحـــل: لنقم أولاً بتعيين قيمة الثابت $\alpha$ حيث نعلم أنَّ:

$$1 = \lim_{\substack{x\to\infty\\ y\to\infty}} F_{X_2}(x,y) = \lim_{\substack{x\to\infty\\ y\to\infty}} \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X_n}(x,y)\; dx\; dy = \alpha\cdot\int_0^1\int_0^1 x\cdot y\; dx\; dy = \frac{\alpha}{4}$$

ومنه ينتج لدينا $\alpha = 4$.

و أمّا لتعيين دالَّة التوزيع $F_{X_2}$ فإنَّنا سنقوم بتجزئة الساحة $\mathbb{R}^2$ إلى خمسة قطاعات كما في الشكل الجانبي (٢٨ ,٥)، فنجد ما يلي:




الشكل (٢٨ ,٥)

أ- قيمة دالَّة التوزيع $F_{X_2}$ عند كل نقطة $(x,y)$ من القطاع (I) تساوي:

$$F_{X_2}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X_2}(t,u)\; dt\; du = 0$$

ب- قيمة دالَّة التوزيع $F_{X_2}$ عند كل نقطة $(x,y)$ من القطاع (II) تساوي:

$$F_{X_2}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X_2}(t,u)\; dt\; du = \int_{0}^{x}\int_{0}^{y} 4\cdot t\; u\;\; dt\, du \; = x^2\cdot y^2$$

ج- قيمة دالَّة التوزيع $F_{X_2}$ عند كل نقطة $(x,y)$ من القطاع (III) تساوي:

$$F_{X_2}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X_2}(t,u)\; dt\; du = \int_{0}^{1}\int_{0}^{y} 4\cdot t\; u\;\; dt\, du \; = y^2$$

د - قيمة دالَّة التوزيع $F_{X_2}$ عند كل نقطة $(x,y)$ من القطاع (IV) تساوي:

$$F_{X_2}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X_2}(t,u)\; dt\; du = \int_{0}^{x}\int_{0}^{1} 4\cdot t\; u\;\; dt\, du \; = x^2$$

هـ- قيمة دالَّة التوزيع $F_{X_2}$ عند كل نقطة $(x,y)$ من القطاع (V) تساوي:

$$F_{X_2}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X_2}(t,u)\; dt\; du = \int_{0}^{1}\int_{0}^{1} 4\cdot t\; u\;\; dt\, du \; = 1$$

ومن ثَّم يكون لدالَّة التوزيع $F_{X_2}$ العرض الآتي:

$$F_{X_2}(x\;,y) = \begin{cases} 0 & for \quad x \le 0 \quad or \quad y \le 0 \\ x^2 y^2 & for \quad 0 < x \le 1 \quad and \quad 0 < y \le 1 \\ y^2 & for \quad x > 1 \quad and \quad 0 < y \le 1 \\ x^2 & for \quad 0 < x \le 1 \quad and \quad y > 1 \\ 1 & for \quad x > 1 \quad and \quad y > 1 \end{cases}$$

وبهذا ينتهي حل المثال المُعطى.

## (٥,٥) التوزيعات الهامشية

**Marginal Distributions**

يصادفنا في كثير من المسائل ضرورة التعامل مع دالَّة التوزيع الاحتمالية لإحدى مركِّبات (أو لأكثر من مركِّبة) متَّجه عشوائي $\mathbb{X}_n = (X_1, X_2, ..., X_n)$، وفي مثل هذه الحالات نكون معنيين باستخلاص صيغة دالَّة التوزيع لهذه المركِّبة (أو المركِّبات) من صيغة دالَّة التوزيع للمتَّجة العشوائي $\mathbb{X}_n$. إنَّ ما يُعرف باسم التوزيعات الهامشيّة تقوم بتنفيذ هذه المهمة.

### (١ , ٥ , ٥) دوال الكثافة الهامشيّة Marginal Density Functions

من أجل توضيح مفهوم التوزيعات الهامشيّة وتبسيطها سنقوم بعرض دوال الكثافة الاحتمالية (أو الكتل الاحتمالية) الهامشيّة لمركِّبة واحدة، ومن ثمَّ لمركِّبتين، وبناءً على تلك الدراسة يمكن استنتاج دوال الكثافة الاحتمالية (أو الكتل الاحتمالية) الهامشيّة لثلاث مركِّبات أو أكثر.

#### (١ , ١ , ٥ , ٥) دالَّة الكثافة الهامشيّة لمركِّبة من مركِّبات متَّجه عشوائي

ليكن $\mathbb{X}_n = (X_1, X_2, ..., X_n)$ متَّجهاً عشوائياً بدالَّة كثافة احتمالية $f_{\mathbb{X}_n}$، علماً أنَّ $n \geq 2$ عدد طبيعي، فعندئذ:

١- إذا كان المتَّجة العشوائي $\mathbb{X}_n$ متقطِّعاً بدالَّة كتلة احتمالية $p_{i_1 i_2 ... i_n}$ فإنَّ دالَّة الكتلة الاحتمالية الهامشيّة للمركبة $X_k$ (علماً أنَّ $k \in \mathbb{N}_n$ كيفي ولكن مثبَّت) تُعرَّف من خلال العلاقة الآتية:

$$f_{X_k}(x_k) \hat{=} p_{i_1 i_2 ... i_{n-1} \bullet i_{k+1} ... i_n} := \sum_{\substack{i_1 \in I_1 \,\&\, i_2 \in I_2 \,\&\cdots\&\, i_{k-1} \in I_{k-1} \\ \&\, i_{k+1} \in I_{k+1} \,\&\cdots\&\, i_n \in I_n}} p_{i_1 i_2 ... i_n} \qquad [5,51]$$

وذلك من أجل أي $x_k$ من مجموعة قيم $X_k$، علماً أنَّ $p_{i_1 i_2 ... i_n}$ مُعطى بالعلاقة [5,42] .

٢- إذا كان المتَّجة العشوائي $\mathbb{X}_n$ مستمراً بدالَّة كثافة احتمالية $f_{\mathbb{X}_n}$ فإنَّ دالَّة الكثافة الهامشيّة للمركِّبة $X_k$ (مع $k \in \mathbb{N}_n$ كيفي ولكن مثبَّت) تُعرَّف من خلال العلاقة الآتية؛ وذلك من أجل كل $x_k \in \mathbb{R}$:

$$f_{X_k}(x_k) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{(n-1)-integrals} f_{\mathbb{X}_n}(x_1, x_2, ..., x_n)\; dx_1 dx_2 ... dx_{k-1} dx_{k+1} ... dx_n \qquad [5,52]$$

#### (٢ , ١ , ٥ , ٥) دالَّة الكثافة الهامشيّة لمركِّبتين من مركِّبات متَّجه عشوائي

ليكن $\mathbb{X}_n = (X_1, X_2, ..., X_n)$ متَّجهاً عشوائياً بدالَّة كثافة احتمالية $f_{\mathbb{X}_2}$، علماً أنَّ $n \geq 2$ عدد طبيعي، ولنأخذ $\ell, k \in \mathbb{N}_n$ مع $k \neq \ell$ كيفيين ولكن مثبَّتين، فعندئذ:

١- إذا كان المتَّجة العشوائي $\mathbb{X}_n$ متقطِّعاً بدالَّة كتلة احتمالية $p_{i_1 i_2 ... i_n}$ فإنَّ الدالَّة الاحتمالية الكتلية الهامشيّة المشتركة للمركبتين $X_\ell$, $X_k$ (علماً أنَّ العددين $\ell, k \in \mathbb{N}_n$ كيفيين ولكنَّهما مثبَّتان مع $k \neq \ell$) تُعرَّف من خلال العلاقة الآتية:

$$f_{X_\ell, X_k}(x_\ell, x_k) \hat{=} p_{i_1 i_2 ... i_{\ell-1} \bullet i_{\ell+1} ... i_{k-1} \bullet i_{k+1} ... i_n} = \sum_{\substack{i_1 \in I_1 \,\&\, i_2 \in I_2 \,\&\cdots\&\, i_{\ell-1} \in I_{\ell-1} \,\&\, i_{\ell+1} \in I_{\ell+1} \\ \&\cdots\&\, i_{k-1} \in I_{k-1} \,\&\, i_{k+1} \in I_{k+1} \,\&\cdots\&\, i_n \in I_n}} p_{i_1 i_2 ... i_n} \qquad [5,53]$$

وذلك من أجل أي $x_k$ و $x_\ell$ من مجموعة قيم $X_k$ و $X_\ell$ على الترتيب.

٢- إذا كان المتَّجة العشوائي $\mathbb{X}_n$ مستمراً بدالَّة كثافة احتمالية $f_{\mathbb{X}_n}$ فإنَّ **دالَّة الكثافة الهامشيّة المشتركة للمركبتين** $X_\ell, X_k$ مع $k \neq \ell$ تُعرَّف من خلال العلاقة الآتية وذلك من أجل كل $x_\ell, x_k \in \mathbb{R}$ :

$$f_{X_\ell, X_k}(x_\ell, x_k) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{(n-2)-integrals} f_{\mathbb{X}_n}(x_1, x_2, \ldots, x_n) dx_1 dx_2 \ldots dx_{\ell-1} dx_{\ell+1} \ldots dx_{k-1} dx_{k+1} \ldots dx_n \qquad [5,53]$$

وعلى هذا النحو تُعمَّم هذه العلاقات على أي عددمن المركبات للمتَّجه $\mathbb{X}_n$، ولكن الصيغ قد لا تبدو بسيطة كما في الحالتين السابقتين.

## (٢ , ٥ , ٥) دوال التوزيع الهامشيّة Marginal Distribution Functions

لننتقل الآن إلى عروض دوال التوزيع الهامشيّة لمتَّجه عشوائي مستخدمين الأسلوب السّابق في التوضيح.

### (١ , ٢ , ٥ , ٥) دوال التوزيع الهامشيّة لمركِّبة من مركِّبات متَّجه عشوائي

ليكن $\mathbb{X}_n = (X_1, X_2, \ldots, X_n)$ متَّجهاً عشوائياً بدالَّة توزيع احتمالية $F_{\mathbb{X}_n}$، فعندئذ:

١- إذا كان المتَّجة العشوائي $\mathbb{X}_n$ متقطِّعاً بدالَّة احتمالية كتلية $p_{i_1 i_2 \ldots i_n}$ فإنَّ **دالَّة التوزيع الهامشيّة للمركِّبة** $X_k$ (علماً أنَّ $k \in \mathbb{N}_n$ كيفي ولكن مثبَّت) تُعرَّف من خلال العلاقة الآتية:

$$F_{X_k}(x) = \sum_{i_k; x_{k i_k} < x} p_{i_1 i_2 \ldots i_{k-1} \bullet i_{k+1} \ldots i_n} \quad ; \forall x \in \mathbb{R} \qquad [5,55]$$

٢- إذا كان المتَّجة العشوائي $\mathbb{X}_n$ مستمراً بدالَّة كثافة احتمالية $f_{\mathbb{X}_n}$ فإنَّ **دالَّة التوزيع الهامشيّة للمركِّبة** $X_k$ تُعرَّف من خلال العلاقة الآتية:

$$F_{X_k}(x) = \int_{-\infty}^{x} f_{X_k}(t) dt \quad ; \forall x \in \mathbb{R} \qquad [5,56]$$

### (٢ , ٢ , ٥ , ٥) دالَّة التوزيع الهامشيّة لمركّبتين من مركِّبات متَّجه عشوائي

ليكن $\mathbb{X}_n = (X_1, X_2, \ldots, X_n)$ متَّجهاً عشوائياً بدالَّة كثافة احتمالية $f_{\mathbb{X}_n}$، علماً أنَّ $n \geq 2$ عدد طبيعي، ولنأخذ العددين $\ell, k \in \mathbb{N}_n$ كيفيين ولكنَّهما مثبَّتان مع $k \neq \ell$ ، فعندئذ:

١- إذا كان المتَّجة العشوائي $\mathbb{X}_n$ متقطِّعاً بدالَّة احتمالية كتلية $p_{i_1 i_2 \ldots i_n}$ فإنَّ **دالَّة التوزيع الهامشيّة المشتركة للمركِّبتين** $X_\ell, X_k$ تُعرَّف لأجل كل $x_1, x_2 \in \mathbb{R}$ من خلال العلاقة الآتية:

$$F_{X_\ell, X_k}(x_1, x_2) = \sum_{i_\ell; x_{\ell i_\ell} < x_1 \,\&\, i_k; x_{k i_k} < x_2} p_{i_1 i_2 \ldots i_{\ell-1} \bullet i_{\ell+1} \ldots i_{k-1} \bullet i_{k+1} \ldots i_n} \qquad [5,57]$$

وذلك من أجل أيّ $x_{\ell i_\ell}$ و $x_{k i_k}$ من مجموعة قيم $X_\ell$ و $X_k$ على الترتيب.

٢- إذا كان المتَّجة العشوائي $\mathbb{X}_n$ مستمراً بدالَّة كثافة احتمالية $f_{\mathbb{X}_n}$ فإنَّ **دالَّة التوزيع الهامشيّة المشتركة للمركِّبتين** $X_\ell, X_k$ مع $k \neq \ell$ تُعرَّف وذلك لأجل كل $x_1, x_2 \in \mathbb{R}$ من خلال العلاقة الآتية:

$$F_{X_\ell, X_k}(x_1, x_2) = \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} f_{X_\ell, X_k}(t_1, t_2) \; dt_1 \, dt_2 \qquad [5,58]$$

و على هذا النحو تُعمَّم هذه العلاقات على أي عددمن المركبات للمتَّجه $\mathbb{X}_n$، ولكن الصيغ قد لا تبدو بسيطة كما في الحالتين

السّابقتين.

**(٣,٢,٥,٥) ملاحظة**

من أجل الحالة الخاصة $n = 2$ وعندما تكون مجموعة قيم المتّجه العشوائي $\mathbb{X}_2 = (X, Y)$ منتهية وقليلة العدد نسبياً، فإنّه يمكن حساب الاحتمالات الهامشيّة من خلال ما يُسمّى بـ **جدول التوزيع الهامشي** لـ $\mathbb{X}_2$، فإذا كانت $\mathbf{X} = \{x_1, x_2, \ldots, x_n\}$ و $\mathbf{Y} = \{y_1, y_2, \ldots, y_m\}$ هي مجموعة قيم المتغيِّر العشوائي $X$ و $Y$ على الترتيب، فعندئذ يكون لجدول التوزيع الهامشي لـ $\mathbb{X}_2$ العرض الآتي:

**الجدول (٥,٦)**

| | | قيم المتغيِّر العشوائي $X$ | | | | Total |
|---|---|---|---|---|---|---|
| | | $x_1$ | $x_2$ | ...... | $x_n$ | |
| قيم المتغيِّر العشوائي $Y$ | $y_1$ | $p_{11}$ | $p_{21}$ | ...... | $p_{n1}$ | $p_{\bullet 1}$ |
| | $y_2$ | $p_{12}$ | $p_{22}$ | ...... | $p_{n2}$ | $p_{\bullet 2}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots \;\vdots\;\vdots$ | $\vdots$ | $\vdots$ |
| | $y_m$ | $p_{1m}$ | $p_{2m}$ | ...... | $p_{nm}$ | $p_{\bullet m}$ |
| Total | | $p_{1\bullet}$ | $p_{2\bullet}$ | ...... | $p_{n\bullet}$ | 1 |

علماً أنّه يجب ترتيب قيم $X$ وكذلك $Y$ في الجدول السّابق بشكل متدرّج من القيمة الأصغر إلى القيمة الأكبر، أي إنّه لدينا في هذا الجدول $x_1 < x_2 < \ldots < x_n$ وكذلك $y_1 < y_2 < \ldots < y_m$.

**(٤,٢,٥,٥) أمثلة**

١- ليكن $\mathbb{X}_2 = (X, Y)$ متّجهاً عشوائياً متقطّعاً بدالّة احتمالية كتلية مُعطاة بالعلاقة الآتية:

$$p_{ij} = \begin{cases} \frac{7}{21} & \text{for } i = 1,2 \;\&\; j = 1 \\ \frac{4}{21} & \text{for } i = 1 \;\&\; j = 2 \\ \frac{3}{21} & \text{for } i = 2 \;\&\; j = 2 \end{cases}$$

ولنقم بتعيين التوزيعات الهامشيّة لكل من $X$ و $Y$ في توزيع المتّجه $\mathbb{X}_2$.

الحـــل: من معطيات المثال نجد الآتي:

$$p_{\bullet 1} = p_{11} + p_{21} = \frac{7}{21} + \frac{7}{21} = \frac{14}{21} \quad \& \quad p_{\bullet 2} = p_{12} + p_{22} = \frac{4}{21} + \frac{3}{21} = \frac{7}{21}$$

$$p_{1\bullet} = p_{11} + p_{12} = \frac{7}{21} + \frac{4}{21} = \frac{11}{21} \quad \& \quad p_{2\bullet} = p_{21} + p_{22} = \frac{7}{21} + \frac{3}{21} = \frac{10}{21}$$

ومن ثمّ ينتج لدينا أنّ لدوال التوزيع الهامشيّة العروض الآتية:

$$F_X(x) = \sum_{i : x_i < x} p_{i\bullet} = \frac{11}{21}\mathbf{I}_{(0,1]}(x) + \mathbf{I}_{(1,\infty)}(x)$$

$$F_Y(y) = \sum_{j : y_j < y} p_{\bullet j} = \frac{14}{21}\mathbf{I}_{(0,1]}(y) + \mathbf{I}_{(1,\infty)}(y)$$

٢- ليكن $\mathbb{X}_2 = (X, Y)$ متَّجهاً عشوائياً بدالَّة كثافة احتمالية مُعطاة على النحو الآتي:

$$f_{\mathbb{X}_2}(x,y) = \begin{cases} 4xy & for\ (x,y) \in [0,1]\times[0,1] \\ 0 & otherwise \end{cases}$$

حيث وجدنا في مثال سابق (المثال /٢/ من (٤ , ٤ , ٥)) أنَّ دالَّة التوزيع لهذا المتَّجه العشوائي هي:

$$F_{\mathbb{X}_2}(x,y) = \begin{cases} 0 & for\ \ x \le 0 & or & y \le 0 \\ x^2y^2 & for\ \ 0 < x \le 1 & and & 0 < y \le 1 \\ y^2 & for\ \ x > 1 & and & 0 < y \le 1 \\ x^2 & for\ \ 0 < x \le 1 & and & y > 1 \\ 1 & for\ \ x > 1 & and & y > 1 \end{cases}$$

عندئذ لتعيين الكثافتين الهامشيتين لهذا التوزيع نجد ما يلي:

أ- إنَّ دالَّة الكثافة الهامشيّة لـ $X$ تُعيَّن بوساطة العلاقة الآتية:

$$f_X(x) = \int_{-\infty}^{\infty} f_{\mathbb{X}_2}(x,y)\,dy = \begin{cases} 2x & for\ \ 0 < x \le 1 \\ 0 & for\ \ x \notin (0,1] \end{cases}$$

ب – إنَّ دالَّة الكثافة الهامشيّة لـ $Y$ فإنَّها تعيَّن بوساطة العلاقة الآتية:

$$f_Y(y) = \int_{-\infty}^{\infty} f_{\mathbb{X}_2}(x,y)\,dx = \begin{cases} 2y & for\ \ 0 < y \le 1 \\ 0 & for\ \ y \notin (0,1] \end{cases}$$

وأمّا لتعيين دالتي التوزيع الهامشيتين لهذا التوزيع فإنّنا نجد بما يلي:

أ- إنَّ دالَّة التوزيع الهامشيّة لـ $X$ تعيَّن بوساطة العلاقة الآتية من أجل كل $x \in \mathbb{R}$:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt = \begin{cases} 0 & for\ \ x \le 0 \\ x^2 & for\ \ 0 < x \le 1 \\ 1 & for\ \ x > 1 \end{cases}$$

ب- إنَّ دالَّة التوزيع الهامشيّة لـ $Y$ فإنَّها تعيَّن بوساطة العلاقة الآتية من أجل كل $y \in \mathbb{R}$:

$$F_Y(y) = \int_{-\infty}^{y} f_Y(t)dt = \begin{cases} 0 & for\ \ y \le 0 \\ y^2 & for\ \ 0 < y \le 1 \\ 1 & for\ \ y > 1 \end{cases}$$

وبهذا يتمّ الحل.

من التطبيقات المهمّة للتوزيعات الهامشيّة ما يُعرَف باسم **التوزيعات المركّبة** Compound Distributions، وهي توزيعات لمتغيِّرات عشوائيَّة لها دوال توز يع تابعة لمَعالم، وهذه المَعالم هي بدورها متغيِّرات عشوائيَّة بتوزيع معيّن أيضاً، ولتوضيح دراسة هذا النوع من التوزيعات لنفترض أنَّ $X$ متغيِّر عشوائي خاضع لتوزيع بمَعْلمة $\xi$، وأنَّ المَعْلمة $\xi$ هي متغيِّر عشوائي خاضع لتوزيع معلوم، فعندئذ لإيجاد توزيع $X$ نأخذ متَّجهاً عشوائياً $\mathbb{X}_2 = (X, \xi)$، ومن ثمَّ نوجد توزيع $X$ من خلال حساب توزيعه الهامشي في توزيع المتَّجه $(X, \lambda)$، والمثال الآتي يوضِّح لنا ذلك.

**(٥ , ٢ , ٥ , ٥) مثال**

ليكن $X$ متغيّراً عشوائياً متقطّعاً ويأخذ قيماً صحيحة غير سالبة من $\mathbb{N}_n^o$ ، ومُمَيَّزاً بالعلاقة الآتية:

$$P(X=k)=\binom{n}{k}p^k(1-p)^{n-k}$$

فعندئذ يُقال عن $X$ إنَّه خاضع للتوزيع الحدّاني (أو ذي الحدين) Binomial Distribution بمَعْلمتين $n$ و $p$ (سنأتي على دراسته بشكلٍ مفصل في الفصل القادم)، ولنفترض أنَّ المَعْلمة $n$ هي متغيِّر عشوائي مستقل عن $p$ وله توزيع مُمَيَّز بالعلاقة الآتية:

$$P(n=\ell)=\frac{\mathbf{e}^{-\lambda}\lambda^{\ell}}{\ell!}$$

علماً أنَّ $\lambda>0$ مثبَّت، فعندئذ يُقال عن $n$ إنَّه خاضع لتوزيع بواسون Poisson Distribution بمَعْلمة $\lambda$ (سنأتي على دراسته بشكل مفصل في الفصل القادم)، ولتعيين توزيع المتغيِّر العشوائي $X$ نأخذ المتَّجه العشوائي $\mathbb{X}_2=(X,n)$ ، ومن ثمَّ نقوم بحساب التوزيع الهامشي لـ $X$، فنجد ما يلي:

$$P(X=k)=P(X=k,n=\mathbb{N})=\sum_{\ell=0}^{\infty}P(X=k,n=\ell)$$

وباستخدام صيغ الاحتمالات الشرطية يصبح لدينا:

$$\begin{aligned}P(X=k)&=\sum_{\ell=o}^{\infty}P(X=k\mid n=\ell)\cdot P(n=\ell)=\sum_{\ell=o}^{\infty}\frac{\mathbf{e}^{-\lambda}\lambda^{\ell}}{\ell!}\binom{\ell}{k}p^k(1-p)^{\ell-k}\\&=\frac{\mathbf{e}^{-\lambda}p^k}{k!}\sum_{\ell=k}^{\infty}\frac{\lambda^{\ell}}{(\ell-k)!}(1-p)^{\ell-k}=\frac{\mathbf{e}^{-\lambda}p^k\lambda^k}{k!}\sum_{\ell=k}^{\infty}\frac{[\lambda(1-p)]^{\ell}}{(\ell-k)!}\\&=\frac{\mathbf{e}^{-\lambda}p^k\lambda^k}{k!}\mathbf{e}^{\lambda(1-p)}=\frac{(\lambda p)^k}{k!}\mathbf{e}^{-\lambda p}\end{aligned}$$

وبما أنَّ الطرف الأيمن من العلاقة الأخيرة هي صيغة لمتغيِّر عشوائي خاضع لتوزيع بواسون بمَعْلمة $\lambda\cdot p$ فإنَّ ذلك يعني أنَّ لـ $X$ توزيع بواسوني بمَعْلمة $\lambda\cdot p$ ، وهذا التوزيع يُدعى التوزيع الحدّاني المُركَّب Compound Binomial Distribution، وكتطبيق على هذا المثال سنفترض أنَّ احتمال ولادة طفل ذكر في بلد ما يساوي 0.517 ، وأنَّ $X$ متغيِّر عشوائي راصد لعدد الأطفال الذكور المولودين في أسرة (من هذا البلد) لديها $n$ طفلاً، فعندئذ يكون لهذا المتغيِّر العشوائي توزيع حدَّاني بمَعْلمتين $n$ و $p=0.517$ ، أي إنَّ احتمال أن يوجد $k$ طفلاً من الذكور في أسرة ما لديها $n$ طفلاً يساوي:

$$P(X=k)=\binom{n}{k}\cdot(0.517)^k(0.483)^{n-k}$$

وبفرض أنَّ $n$ متغيِّر عشوائي بواسوني بمَعْلمة $\lambda>0$ ، فعندئذ وفقاً لما توصلنا إليه في المثال السّابق يكون لدينا:

$$P(X=k)=\frac{(\lambda\cdot p)^k}{k!}\mathbf{e}^{-\lambda\cdot p}=\frac{(0.517\lambda)^k}{k!}\mathbf{e}^{-0.517\lambda}$$

وهنا نلاحظ أنَّ هذه العلاقة الأخيرة تتعلّق بقيمة المَعْلمة $\lambda$ التي يُفترض أن تكون معلومة من أجل بلد ما، علماً أنَّ قيمة هذه المَعْلمة تتغيَّر من بلد لآخر.

## (٥،٦) التوزيعات الشرطية

**Conditional Distributions**

من الدراسات المهمَّة في الاحتمالات تلك التوزيعات التي تُعرف باسم التوزيعات الشرطية، حيث تقوم هذه التوزيعات بتعيين توزيع مركّبة أو أكثر من مركّبات متّجه عشوائي مشروطةً بتحقُّق بعض المركّبات المتبقية أو كلها.

### (٥ ، ٦ ، ١) الاحتمالات الشرطية للمتغيِّرات العشوائي

لقد قدّمنا في فصل سابق مفهوم الاحتمال الشرطي لحادث $A$ بالنسبة إلى حادث آخر $B$ له احتمال موجب، ورمزنا لذلك بـ $P(A \mid B)$، والآن سنقدِّم تمهيداً لمفهوم الاحتمال الشرطي لمتغيِّر عشوائي $X$ بالنسبة إلى متغيِّر عشوائي آخر $Y$ من خلال تقديم مفهوم الاحتمال الشرطي لمتغيِّر عشوائي $X$ بالنسبة إلى حادث $B \in \mathscr{A}$ علماً أنَّ $0 < P(B)$. إنَّ هذا المفهوم لا يختلف كثيراً عن مفهوم الاحتمال الشرطي $P(A \mid B)$ سوى أنَّ الحادث الذي سيحلّ مكان $A$ هو الحادث $\{\omega \in \Omega \ ; X(\omega) \in D\}$ مع $D \in \Re$.

#### (٥ ، ٦ ، ١ ، ١) الاحتمال الشرطي لمغيِّر عشوائي بالنسبة إلى حادث

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ ولنأخذ $B \in \mathscr{A}$ مع $0 < P(B)$. عندئذ يُعرَّف **قانون التوزيع الشرطي** من $X$ بالنسبة إلى (أو تحت شرط تحقُّق) الحادث $B$ من خلال العلاقة الآتية:

$$P_{X \mid B}(D) \triangleq P\left(X \in D \mid B\right) = P_B\left(X^{-1}(D)\right) = \frac{P\left(X^{-1}(D) \cap B\right)}{P(B)} \qquad [5,58]$$

وهذا يعني أنَّ $P_{X \mid B} = P_B \circ X^{-1}$ (مبرِّر صحة هذا الاستخدام تجده في الملاحظة / ٣/ من (١-١-٢-٤)). لاحظ هنا أنَّه في الحالة الخاصّة عندما يكون لدينا $D = (-\infty, x)$ مع $x \in \mathbb{R}$ فإنَّنا سنحصل على دالّة التوزيع الشرطية من $X$ بالنسبة إلى الحادث $B$ التي تُكتب بالشكل $F_{X \mid B}$ حيث لدينا:

$$\begin{aligned} F_{X \mid B}(x) &= P_{X \mid B}\left((-\infty, x)\right) = P_B\left(X^{-1}(-\infty, x)\right) \\ &= P\left(X \in (-\infty, x) \mid B\right) = P\left(X < x \mid B\right) = F_X\left(x \mid B\right) \end{aligned}$$

#### (٥ ، ٦ ، ١ ، ٢) الاحتمال الشرطي لحادث بالنسبة إلى متغيِّر عشوائي

لنوضِّح الآن مفهوم الاحتمال الشرطي لحادث $B$ بالنسبة إلى متغيِّر عشوائي $X$ (أو تحت شرط تحقُّق الحادث $X \in D$) حيث سنفترض على سبيل التبسيط والتوضيح أنَّ هذا المتغيِّر العشوائي متقطِّع.

لتكن $X$ متغيِّراً عشوائياً متقطِّعاً بمجموعة قيم $\mathbf{X} := \{x_i \ ; i \in I\}$ ومعرَّف فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، ولنضع $Z_i := X^{-1}(\{x_i\})$ من أجل أيّة قيمة $i \in I$، فنجد أنَّ $Z_i \neq \varnothing$ من أجل كل $i \in I$، وأنَّ الصف $\mathscr{Z}_X := \{Z_i \ ; i \in I\}$ يشكِّل تجزئة للحادث الأكيد $\Omega$ (ولكنَّه لا يشكِّل جبراً فوق $\Omega$)، ويمكننا أن نولِّد منه $\sigma$ – **جبر** فوق $\Omega$ من خلال أخذ صف كل الاجتماعات المنتهية لعناصر من هذا الصف الذي يُعرف باسم الـ $\sigma$ – **جبر المولَّد بالمتغيِّر العشوائي $X$**، فلو رمزنا لهذا الـ $\sigma$ – **جبر** بالرمز $\mathscr{A}_X$ فإنَّه سيكون لـ $\mathscr{A}_X$ العرض الآتي:

$$\mathscr{A}_X := \left\{ \bigcup_{k \in K} Z_k \ ; K \subseteq I \text{ with } K \text{ finite} \right\} \qquad [5,59]$$

فإذا أخذنا $\mathscr{A} = 2^{\Omega}$ فعندئذ سيكون $\mathscr{A} \supseteq \mathscr{A}_X$.

بعد هذا التقديم يمكننا تقديم تعريف الاحتمال الشرطي لحادث $B$ من $\mathscr{A}$ بالنسبة إلى متغيِّر عشوائي $X$ على أنَّه **متغيِّر**

عشوائي فوق الفضاء الاحتمالي $[\Omega, \mathscr{A}, P]$ نفسه، ومُعطى على النحو الآتي:

$$P(B \mid X)(\omega) = \begin{cases} \dfrac{P(B \cap Z_i)}{P(Z_i)} & for \ \ \omega \in Z_i \ with \ P(Z_i) > 0 \\ 0 \leq any \ value \leq 1 & for \ \ \omega \in Z_i \ with \ P(Z_i) = 0 \end{cases} \qquad [5,60]$$

وفي هذه الحالة سيكون المتغيِّر العشوائي $P(B \mid X)(\omega)$ متقطِّعاً أيضاً لأنَّ عدد عناصر التجزئة $\mathscr{Z}$ قابل للعدِّ على الأكثر، وكذلك يُلاحظ أنَّه من أجل كل $A \in \mathscr{A}_X$ تكون العلاقة الآتية محقَّقة:

$$P(B \mid A) = \sum_{\substack{k \in K \ ; Z_{i_k} \subseteq A \\ P(Z_{i_k}) > 0}} P\left(B \cap Z_{i_k}\right) \cdot P\left(Z_{i_k}\right)$$

ويقابلها من أجل المتغيِّرات العشوائيّة المستمرَّة العلاقة الآتية:

$$P(B \mid A) = \int_A P(B \cap X)(\omega) \, P(d\omega) \qquad [5,61\text{-a}]$$

وفي الحالة الخاصّة عندما نأخذ $A = \Omega$ ينتج لدينا العلاقة الآتية:

$$P(B) = \int_\Omega P(B \cap X)(\omega) \, P(d\omega) \qquad [5,61\text{-b}]$$

وهذه العلاقة الأخيرة تُعدّ تعميماً لصيغة الاحتمال التام التي مرَّت معنا في الفصل السّابق.

إنَّ دراسة ا لتوزيعات الشرطية للمتَّجهات العشوائيّة لا تخلو من الصعوبة عند التعامل مع متَّجهات عشوائية ذات بُعد أكبر من ثلاث (3)، وهذه الصعوبة تكمن بالدرجة الأولى في شكل الصياغات المستخدمة ودقّة التعامل معها، وهذا ما سنلاحظه لاحقاً أثناء الدراسة في الفقرة القادمة، ولذلك سنقوم أولاً بتقديم الدراسة من أجل متَّجه عشوائي ثنائي البُعد ومن ثَّم تعميم ذلك على متَّجه عشوائي ثلاثي البُعد للحالة الموافقة للمتَّجهات العشوائيّة المتقطِّعة والمستمرَّة فقط، ولن نقوم بعرض الحالة الموافقة للمتَّجهات العشوائيّة المتقطِّعة ذات البُعد النوني حيث يمكن لمن يود الاطلاع عليها الرجوع إلى المراجع ذات الصلة ومنها [64] المذكور في فهرس المراجع في نهاية هذا الكتاب.

إنَّ فكرة التوزيعات الشرطية مبني على فهم المعنى للاحتمال الشرطي لمتغيِّر عشوائي $X$ بالنسبة (تحت شرط تحقُّق حادث متعلِّق بمتغيِّر عشوائي آخر $Y$) إلى متغيِّر عشوائي $Y$ ولذلك سنقوم أولاً بتوضيح معنى الاحتمال الشرطي لمتغيِّر عشوائي $X$ النسبة إلى متغيِّر عشوائي $Y$، وهنا سنفرِّق بين الحالات التي يكون فيها المتَّجه العشوائي $\mathbb{X}_2 = (X \ , \ Y)$ متقطِّعاً أو مستمرًّا وذلك على سبيل التوضيح والتبسيط في الدراسة، ومن ثَّم سنعمِّم التعريف على الحالة $\mathbb{X}_3 = (X_1, X_2, X_3)$ لأنَّ عملية التعميم تتمّ بطريقة مماثلة.

## (٥،٦،٢) التوزيعات الشرطية لمتَّجه عشوائي متقطِّع Conditional Distributions of a Discrete Random Vector

سنعالج فيما يلي الحالة الموافقة للمتَّجهات العشوائيّة المتقطِّعة فقط، وسنبدأها بالدوال الاحتمالية الشرطية لمتَّجه عشوائي متقطِّع.

### (٥،٦،٢،١) الدوال الاحتمالية الشرطية لمتَّجه عشوائي متقطِّع

ليكن $\mathbb{X}_2 = (X \ , \ Y)$ متَّجهاً عشوائياً متقطِّعاً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، ولنفترض أنَّ $\mathbf{X} = \{x_i \ ; \ i \in I \subseteq \mathbb{N}\}$ و $\mathbf{Y} = \{y_j \ ; \ j \in J \subseteq \mathbb{N}\}$ هي مجموعة قيم $X$ و $Y$ على الترتيب، فعندئذ دالَّة الاحتمال (أو الكثافة الاحتمالية) الشرطية للمتغيِّر العشوائي $X$ عندما يأخذ القيمة $x_i$ علماً أنَّ المتغيِّر العشوائي $Y$ قد أخذ القيمة $y_j$ باحتمال موجب تماماً تُعطى بالعلاقة الآتية:

$$f_{X|Y}(x_i \mid y_j) = P\left(X = x_i \mid Y = y_j\right) = \frac{P(X = x_i \ \wedge Y = y_j)}{P(Y = y_j)}$$

فإذا رمزنا بـ $p_{i|j} \triangleq P\left(X = x_i \mid Y = y_j\right)$، ونعلم أنَّ:

$$P(Y = y_j) = \sum_{i \in I} P\left(X = x_i \wedge Y = y_j\right) = p_{\bullet j}$$

فإنَّه يمكننا أن نكتب:

$$p_{i|j} = \frac{p_{ij}}{p_{\bullet j}} \qquad [5,62]$$

وبشكل مماثل نجد أنَّ دالَّة الاحتمال الشرطية للمتغيِّر العشوائي $Y$ عندما يأخذ القيمة $y_j$ علماً أنَّ $X$ قد أخذ القيمة $x_i$ باحتمال موجب تماماً هي:

$$f_{Y|X}(y_j \mid x_i) = P\left(Y = y_j \mid X = x_i\right) = \frac{P(X = x_i \ \wedge Y = y_j)}{P(X = x_i)}$$

ولدينا:

$$P(X = x_i) = \sum_{j \in J} P\left(X = x_i \wedge Y = y_j\right) = p_{j\bullet}$$

فإنَّه يمكننا أن نكتب:

$$p_{j|i} = \frac{p_{ji}}{p_{j\bullet}} \qquad [5,63]$$

الآن لنأخذ $\mathbb{X}_3 = (X, Y, Z)$ متَّجهًا عشوائيًّا متقطِّعًا فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، ولنفترض أنَّ:

$$\mathbf{X} = \left\{x_i \,;\, i \in I \subseteq \mathbb{N}\right\} \quad \& \quad \mathbf{Y} = \left\{y_j \,;\, j \in J \subseteq \mathbb{N}\right\} \quad \& \quad \mathbf{Z} = \left\{z_k \,;\, k \in K \subseteq \mathbb{N}\right\}$$

هي مجموعة قيم المتغيِّرات العشوائية $X$ و$Y$ و$Z$ على الترتيب، فعندئذ دالَّة الاحتمال الشرطية للمتغيِّر العشوائي $X$ عندما يأخذ القيمة $x_i$ علماً أنَّ المتَّجه $(Y, Z)$ قد أخذ القيمة $(y_j, z_k) \in \mathbf{Y} \times \mathbf{Z}$ باحتمال موجب تماماً تُعطى بالعلاقة الآتية:

$$f_{X|Y\,Z}(x_i \mid y_j, z_k) = P(X = x_i \mid Y = y_j, Z = z_k) = \frac{P(X = x_i \wedge Y = y_j \wedge Z = z_k)}{P(Y = y_j \wedge Z = z_k)}$$

فلو رمزنا لـ $f_{X|Y\,Z}(x_i \mid y_j, z_k)$ بالرمز $p_{i|jk}$، وكذلك وضعنا:

$$P(Y = y_j \wedge Z = z_k) = \sum_{i \in I} P(X = x_i, Y = y_j, Z = z_k) = p_{\bullet jk}$$

فإنَّه يمكننا أن نكتب:

$$p_{i|jk} = \frac{p_{ijk}}{p_{\bullet jk}} \qquad [5,64\text{-a}]$$

وبشكل مماثل فلو رمزنا لـ $f_{Y|X\,Z}(y_j \mid x_i, z_k)$ و $f_{Z|X\,Y}(z_k \mid x_i, y_j)$ بالرمز $p_{j|ik}$ و $p_{k|ij}$ على الترتيب، وكذلك وضعنا:

$$P(X = x_i, Z = z_k) = \sum_{j \in J} P(Y = y_j, X = x_i, Z = z_k) = p_{i\bullet k}$$

$$P(X=x_i, Y=y_j) = \sum_{k\in K} P(Z=z_k, X=x_i, Y=y_j) = p_{ij\bullet}$$

فإنَّه يمكننا أن نكتب:

$$p_{j|ik} = \frac{p_{ijk}}{p_{i\bullet k}} \quad [5,64\text{-b}]$$

$$p_{k|ij} = \frac{p_{ijk}}{p_{ij\bullet}} \quad [5,64\text{-c}]$$

وأمّا دالَّة الاحتمال الشرطية للمتَّجه $(X, Y)$ عندما يأخـذ القيمة $(x_i, y_j) \in \mathbf{X}\times\mathbf{Y}$ علماً أنَّ المتغيِّر العشوائي $Z$ قد أخذ القيمة $z_k$ باحتمال موجب تماماً، فإنَّها تُعطى بالعلاقة الآتية:

$$f_{XY|Z}\left(x_i, y_j \mid z_k\right) = P\left(X=x_i, Y=y_j \mid Z=z_k\right) = \frac{P(X=x_i \wedge Y=y_j \wedge Z=z_k)}{P(Z=z_k)}$$

فلو رمزنا لـ $f_{XY|Z}(x_i, y_j \mid z_k)$ بالرمز $p_{ij|k}$، وكذلك وضعنا:

$$P(Z=z_k) = \sum_{i\in I}\sum_{j\in J} P(X=x_i, Y=y_j, Z=z_k) = p_{\bullet\bullet k}$$

فإنَّه يمكننا أن نكتب:

$$p_{ij|k} = \frac{p_{ijk}}{p_{\bullet\bullet k}} \quad [5,65\text{-a}]$$

وبشكل مماثل فلو رمزنا لـ $f_{XZ|Y}(x_i, z_k \mid y_j)$ و $f_{YZ|X}(y_j, z_k \mid x_i)$ بالرمز $p_{ik|j}$ و $p_{jk|i}$ على الترتيب، ووضعنا بالتعريف ما يلي:

$$P(Y=x_j) = \sum_{i\in I}\sum_{k\in K} P(X=x_i, Z=z_k, Y=x_j) = p_{\bullet j\bullet}$$

$$P(X=x_i) = \sum_{j\in J}\sum_{k\in K} P(Y=y_j, Z=z_k, X=x_i) = p_{i\bullet\bullet}$$

فإنَّه يمكننا أن نكتب الآتي:

$$p_{ik|j} = \frac{p_{ijk}}{p_{\bullet j\bullet}} \quad [5,65\text{-b}]$$

$$p_{jk|i} = \frac{p_{ijk}}{p_{i\bullet\bullet}} \quad [5,65\text{-c}]$$

إنَّ الصيغ الموافقة للحالة التي يكون فيها $\mathbb{X}_n = (X_1, X_2, ..., X_n)$ متَّجهاً عشوائياً متقطِّعاً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ تعمَّم بأسلوب مماثل، ولكنَّ عمليّة الترميز، وصياغة العلاقات لن تكون بسيطة كما هو الحال لدى الحالتين السّابقتين.

### (٥,٦,٢,٢) دالَّة التوزيع الشرطية لمتَّجه عشوائي متقطِّع

لنأخذ متَّجه عشوائي متقطِّع $\mathbb{X}_3 = (X, Y, Z)$ فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، ولنفترض أنَّ $\mathbf{X} = \{x_i ; i\in I \subseteq \mathbb{N}\}$ و $\mathbf{Y} = \{y_j ; j\in J \subseteq \mathbb{N}\}$ و $\mathbf{Z} = \{z_k ; k\in K \subseteq \mathbb{N}\}$ هي مجموعة قيم $X$ و $Y$ و $Z$ على الترتيب، فعندئذ دالَّة التوزيع الشرطية للمتغيِّر العشوائي $X$ عندما يأخذ المتَّجه $(Y, Z)$ قد أخذ القيمة $(y_j, z_k) \in \mathbf{Y}\times\mathbf{Z}$ باحتمال موجب تماماً

تُعطى بالعلاقة الآتية:

$$F_{X|Y\,Z}(x \mid y_j, z_k) = P(X < x \mid Y = y_j, Z = z_k) = \sum_{i \in I\,;\, x_i < x} p_{i\,|\,j\,k} \qquad [5,66\text{-a}]$$

وبالمثل نجد أنَّ دالَّة التوزيع الشرطية للمتغيِّر العشـوائي $Y$ عندما يأخـذ المتَّجه $(X\ ,\ Z)$ القيمة $(x_i, z_k) \in \mathbf{X}\times\mathbf{Z}$ باحتمال موجب تماماً تُعطى بالعلاقة الآتية:

$$F_{Y|X\,Z}(y \mid x_i, z_k) = P(Y < y \mid X = x_i, Z = z_k) = \sum_{j \in J\,;\, y_j < y} p_{j\,|\,i\,k} \qquad [5,66\text{-b}]$$

وكذلك دالَّة التوزيع الشرطية للمتغيِّر العشـوائي $Z$ عندما يأخـذ المتَّجه $(X\ ,\ Y)$ القيمة $(x_i, y_j) \in \mathbf{X}\times\mathbf{Y}$ باحتمال موجب تماماً تُعطى بالعلاقة الآتية:

$$F_{Z|X\,Y}(z \mid x_i, y_j) = P(Z < z \mid X = x_i, Y = y_j) = \sum_{k \in K\,;\, z_k < z} p_{k\,|\,i\,j} \qquad [5,66\text{-c}]$$

أمّا دالَّة التوزيع الشرطية للمتَّجه $(X\ ,\ Y)$ عندما يأخذ المتغيِّر العشوائي $Z$ القيمة $z_k$ باحتمال موجب تماماً فإنَّها تُعطى من خلال العلاقة الآتية وذلك من أجل كل $(x, y) \in \mathbb{R}^2$:

$$F_{X\,Y|Z}(x, y \mid z_k) = P(X < x, Y < y \mid Z = z_k) = \sum_{\substack{i \in I\,;\, x_i < x \\ j \in J\,;\, y_j < y}} p_{i\,j\,|\,k} \qquad [5,67\text{-a}]$$

وبالمثل نجد أنَّ دالَّة التوزيـع الشـرطية للمتَّجـه $(X\ ,\ Z)$ عندما يأخذ المتغيِّر العشوائي $Y$ القيمة $y_j$ باحتمال موجب تماماً تُعطى من خلال العلاقة الآتية، وذلك من أجل كل $(x, z) \in \mathbb{R}^2$:

$$F_{X\,Z|Y}(x, z \mid y_j) = P(X < x, Z < z \mid Y = y_j) = \sum_{\substack{i \in I\,;\, x_i < x \\ k \in K\,;\, z_k < z}} p_{i\,k\,|\,j} \qquad [5,66\text{-b}]$$

وكذلك دالَّة التوزيع الشرطية للمتَّجه $(Y\ ,\ Z)$ عندما يأخذ المتغيِّر العشوائي $X$ القيمة $x_i$ باحتمال موجب تماماً فإنَّها تُعطى من خلال العلاقة الآتية؛ وذلك من أجل كل $(y, z) \in \mathbb{R}^2$:

$$F_{Y\,Z|X}(y, z \mid x_i) = P(Y < y, Z < z \mid X = x_i) = \sum_{\substack{j \in J\,;\, y_j < y \\ k \in K\,;\, z_k < z}} p_{j\,k\,|\,i} \qquad [5,66\text{-c}]$$

وأخيراً نجد أنَّ دالَّة التوزيع الشرطية للمتغيِّر العشوائي $X$ علماً أنَّ $Y$ قد أخذ القيمة $y_j \in \mathbf{Y}$ و $Z \in \mathbb{R}$ باحتمال موجب تماماً تُعطى من خلال العلاقة الآتية وذلك من أجل كل $x \in \mathbb{R}$:

$$F_{X|Y\,Z}(x \mid y_j, \infty) = P(X < x \mid Y = y_j, Z \in \mathbb{R})$$

وبملاحظة أنَّ الطرف الأيمن من العلاقة يمكن كتابته على الشكل الآتي:

$$P(X < x \mid Y = y_j, Z \in \mathbb{R}) = \sum_{k \in K} P(X < x \mid Y = y_j, Z = z_k)$$

فإنَّه يمكننا أن نكتب:

$$F_{X|Y\,Z}(x \mid y_j, \infty) = \sum_{i \in I\,;\, x_i < x} p_{i\,|\,j\bullet} \qquad [5,67\text{-a}]$$

وتجدر الإشارة هنا إلى أنَّ الطرف الأيسر من العلاقة الأخيرة هي كتابة اصطلاحية، ويجب ألا يفهم منها أنَّ $z_k = \infty$، وإنَّما المقصود

منها هو أنَّ $Z(\omega) \in \mathbb{R}$ من أجل كل $\omega \in \Omega$ .

بالمثل نجد أنَّ دالَّة التوزيع الشرطية للمتغيِّر العشوائي $Y$ علماً أنَّ $X$ قد أخذ القيمة $x_i \in \mathbf{X}$ و $Z \in \mathbb{R}$ باحتمال موجب تماماً تُعطى من خلال العلاقة الآتية؛ وذلك من أجل كل $y \in \mathbb{R}$ :

$$F_{Y|XZ}(y \mid x_i, \infty) = P(Y < y \mid X = x_i, Z \in \mathbb{R}) = \sum_{j \in J;\, y_j < y} p_{j|i\bullet} \qquad [5,67\text{-b}]$$

وكذلك نجد أنَّ دالَّة التوزيع الشرطية للمتغيِّر العشوائي $Z$ علماً أنَّ $X$ قد أخذ القيمة $x_i \in \mathbf{X}$ و $Y \in \mathbb{R}$ باحتمال موجب تماماً تُعطى من خلال العلاقة الآتية؛ وذلك من أجل كل $z \in \mathbb{R}$ :

$$F_{Z|XY}(z \mid x_i, \infty) = P(Z < z \mid X = x_i, Y \in \mathbb{R}) = \sum_{k \in K;\, z_k < z} p_{k|i\bullet} \qquad [5,67\text{-c}]$$

وعلى هذا المنوال تُناقش بقية الحالات الممكنة بين $X$ و $Y$ و $Z$.

**(٣،٢،٦،٥) مثال**

لدينا صندوق يحوي 36 كرة متماثلة تماماً، ومرقَّمة بالأرقام 1 وحتى 36. نخلط الكرات جيداً، ومن ثَّم نسحب كرة من الصندوق دون إعادتها، وسنأخذ متَّجهاً عشوائياً $\mathbb{X}_2 = (X, Y)$ فوق الفضاء الاحتمالي لهذه التجربة العشوائيّة بحيث إنَّ:

$$X(\omega) = \begin{cases} 0 & \text{for } \omega \text{ even} \\ 1 & \text{for } \omega \text{ odd} \end{cases} \qquad \& \qquad Y(\omega) = \begin{cases} 0 & \text{for } \omega \text{ divisible on } 3 \\ 1 & \text{otherwise} \end{cases}$$

ولنقم بتعيين التوزيعات الهامشيّة والشرطية لمركبات هذا المتَّجه العشوائي.

الحل: لنقم أولاً بتعيين التوزيعات الهامشيّة لكل من $X$ و$Y$ في توزيع $\mathbb{X}_2 = (X, Y)$ ، فنجد أنَّ مجموعة قيم $X$ هي $\mathbf{X} = \{x_1 = 0, x_2 = 1\}$ وكذلك مجموعة قيم $Y$ هي $\mathbf{Y} = \{y_1 = 0, y_2 = 1\}$ ، ومن ثَّم يكون لدينا ما يلي:

$$p_{11} = P(X = x_1, Y = y_1) = P(\{6, 12, 18, 24, 30, 36\}) = \frac{6}{36}$$

$$p_{12} = P(X = x_1, Y = y_2) = P(\{2, 4, 8, 10, 14, 16, 20, 22, 26, 28, 32, 34\}) = \frac{12}{36}$$

$$p_{21} = P(X = x_2, Y = y_1) = P(\{3, 9, 15, 21, 27, 33\}) = \frac{6}{36}$$

$$p_{22} = P(X = x_2, Y = y_2) = P(\{1, 5, 7, 11, 13, 17, 19, 23, 25, 29, 31, 35\}) = \frac{12}{36}$$

وأمَّا التوزيعات الهامشيّة لهذا المتَّجه العشوائي فنجدها تساوي:

$$p_{1\bullet} = p_{11} + p_{12} = \frac{6}{36} + \frac{12}{36} = \frac{18}{36}$$

$$p_{2\bullet} = p_{21} + p_{22} = \frac{6}{36} + \frac{12}{36} = \frac{18}{36}$$

$$p_{\bullet 1} = p_{11} + p_{21} = \frac{6}{36} + \frac{6}{36} = \frac{12}{36}$$

$$p_{\bullet 2} = p_{12} + p_{22} = \frac{12}{36} + \frac{12}{36} = \frac{24}{36}$$

ومن ثَّم ينتج لدينا ما يلي:

$$F_X(x) = \sum_{i\ ; x_i < x} p_{i \bullet} = \begin{cases} 0 & for\ x \le 0 \\ p_{1\bullet} = \dfrac{1}{2} & for\ 0 < x \le 1 \\ p_{1\bullet} + p_{2\bullet} = 1 & for\ x > 1 \end{cases}$$

وبالمثل نجد:

$$F_Y(y) = \sum_{j\ ; y_j < y} p_{\bullet j} = \begin{cases} 0 & for\ y \le 0 \\ p_{\bullet 1} = \dfrac{1}{3} & for\ 0 < y \le 1 \\ p_{\bullet 1} + p_{\bullet 2} = 1 & for\ y > 1 \end{cases}$$

ولتعيين دالَّة التوزيـع الشـرطية لـ $X$ عندما يكون الحادث $\{Y = y_2\}$ مُحقَّقاً (أي تعيين $F_{X|Y}(x \mid y_2)$ ) نلاحظ أنَّ:

$$P(Y = y_2) = \frac{24}{36} > 0$$

ومنه يكون لدينا:

$$F_{X|Y}(x \mid y_2) = \sum_{i\ ; x_i < x} p_{i|2} = \sum_{i\ ; x_i < x} \frac{p_{i2}}{p_{\bullet 2}} = \frac{1}{p_{\bullet 2}} \sum_{i\ ; x_i < x} p_{i2} = \frac{36}{24} \sum_{i\ ; x_i < x} p_{i2}$$

ومن العلاقة السابقة يكون لدينا:

$$F_{X|Y}(x \mid y_2) = \begin{cases} 0 & for\ x \le 0 \\ \dfrac{36}{24}\dfrac{12}{36} = \dfrac{1}{2} & for\ 0 < x \le 1 \\ 1 & for\ x > 1 \end{cases}$$

و أمَّا تعيين بقية التوزيعات الشرطية لـ $X$ و $Y$ فنترك كتمرين للقارئ؛ لأنَّها تحسب بطريقة مماثلة لما سبق.

### (٣, ٦, ٥) التوزيعات الشرطية لمتَّجه عشوائي مستمر Coditional Distributions of a Continuous Random Vector

في الواقع إنَّ دراسة الكثافات الشرطية للمتَّجهات العشوائيّة المستمرَّة ليست بسيطة كما هو الحال للمتَّجهات العشوائيّة المتقطّعة؛ لأنَّ ذلك ينطوي على شروط عديدة ستتعرَّف عليها في أثناء دراسة هذه الفقرة، ولكن سنقوم أولاً بتوضيح مفهوم الاحتمال الشرطي لمتغيِّر عشوائي مستمرّ $X$ بالنسبة إلى متغيِّر عشوائي مستمرّ $Y$ عندما يأخذ $Y$ قيمة محدَّدة $y$.

نعلم أنَّ احتمال أن يأخذ متغيِّر عشوائي مستمرّ قيمة مُحدّدة تماماً يساوي الصفر دوماً، ومن ثَّم يجب علينا توضيح معنى $P(X < x \mid Y = y)$ من أجل أي $x \in \mathbb{R}$ . من أجل ذلك لنأخذ $\mathbb{X}_2 = (X\ ,\ Y)$ متَّجهاً عشوائياً مستمرّاً بدالَّة كثافة احتمالية $f_{\mathbb{X}_2}(x\ ,\ y)$، ولنفترض أنَّه من أجل أي عدد حقيقي $h > 0$ وكل $y \in \mathbb{R}$ لدينا الحادث:

$$A := \{\, \omega \in \Omega\ \ ;\ y \le Y(\omega) < y + dy \,\} \qquad [5,68]$$

مُحقَّقاً باحتمال موجب تماماً، أي إنَّ $0 < P(A)$ ، فعندئذ يمكننا أن نكتب الآتي:

$$P\left(X<x \mid y \leq Y<y+h\right)=\frac{P\left(X<x \wedge y \leq Y<y+h\right)}{P\left(y \leq Y<y+h\right)}=\frac{\int_{y}^{y+h}\int_{-\infty}^{x} f_{\mathbb{X}_2}(t,u)\ dt\ du}{\int_{y}^{y+h}\int_{-\infty}^{+\infty} f_{\mathbb{X}_2}(t,u)\ dt\ du} \qquad [5,69]$$

فنلاحظ أنَّ الاحتمال في العلاقة الأخيرة ما هو إلاَّ دالَّة التوزيع الشرطي لـ $X$ تحت شرط تحقُّق الحادث $A$ المُعطى بالعلاقة[5,68]، وعلاوة على ذلك فإنَّ الطرف الأيمن من العلاقة الأخيرة هي دالَّة مستمرَّة وغير متناقصة بـ $x$ وتأخذ القيمة 0 عندما $x \to -\infty$ والقيمة 1 عندما $x \to +\infty$. لاحظ أنَّه حتى الآن ما زالت الصعوبة تكمن في إعطاء الاحتمال $P\left(X<x \mid Y=y\right)$ معنى مُحدَّد؛ ولذلك سنقتصر في دراستنا هنا على الحالة التي تكون فيها النهاية $\lim\limits_{h\to 0} P\left(X<x \mid y \leq Y<y+h\right)$ موجودةً، وكذلك سنفترض أنَّ دالَّة الكثافة الهامشيّة $f_Y(y)=\int_{-\infty}^{+\infty} f_{\mathbb{X}_2}(x,y)\,dx$ مستمرَّةً وموجبة تماماً عند كل نقطة $y \in \mathbb{R}$، أي إنَّ $0<f_Y(y)$.

الآن بالعودة إلى العلاقة [5,69] فإنَّنا نجد بتقسيم بسط ومقام الطرف الأيمن من تلك العلاقة على $h$ ومن ثمَّ أخذ نهاية الطرفين عـندما $h \to 0$ فيكون لدينا:

$$\lim_{h\to 0} P\left(X<x \mid y \leq Y<y+h\right)=\lim_{h\to 0}\frac{\frac{1}{h}\int_{y}^{y+h}\int_{-\infty}^{x} f_{\mathbb{X}_2}(t,u)\,dt\,du}{\frac{1}{h}\int_{y}^{y+h}\int_{-\infty}^{+\infty} f_{\mathbb{X}_2}(t,u)\,dt\,du}=\frac{\int_{-\infty}^{x} f_{\mathbb{X}_2}(t,y)dt}{f_Y(y)}$$

ومن ثمَّ يمكننا أن نكتب:

$$P\left(X<x \mid y \leq Y<y+h\right)\approx\frac{\int_{t=-\infty}^{x} f_{\mathbb{X}_2}(t,y)\ dt\ dy}{f_Y(y)\ dy}=\int_{t=-\infty}^{x}\frac{f_{\mathbb{X}_2}(t,y)}{f_Y(y)}\ dt$$

فإذا رمزنا للنهاية $\lim\limits_{h\to 0} P\left(X<x \mid y \leq Y<y+h\right)$ بـ $F_{X\mid Y}(x \mid y)$، فعندئذ سيكون لدينا:

$$F_{X\mid Y}(x \mid y)=\frac{\int_{-\infty}^{x} f_{\mathbb{X}_2}(t,y)\,dt}{f_Y(y)} \qquad [5,70]$$

ومنه ينتج (بعد الأخذ بالحسبان الفرضيات التي قُدِّمت آنفًا) أنَّ $F_{X\mid Y}(x \mid y)$ هي دالَّة توزيع شرطية لمتغيِّر عشوائي مستمر، ولو رمزنا لدالَّة الكثافة الاحتمالية الموافقة للدالَّة $F_{X\mid Y}(x \mid y)$ بالرمز $f_{X\mid Y}(x \mid y)$، فعندئذ يصبح لدينا:

$$f_{X\mid Y}(x \mid y)=\frac{f_{\mathbb{X}_2}(x,y)}{f_Y(y)} \quad ;\forall x \in \mathbb{R} \qquad [5,71\text{-a}]$$

علماً أنَّ $f_Y(y)$ هي دالَّة الكثافة الهامشيّة لـ $Y$، ومن ثمَّ يمكننا أن نكتب الآتي:

$$f_Y(y)\cdot F_{X\mid Y}\left(x \mid y\right)=\int_{-\infty}^{x} f_{\mathbb{X}_2}(x,y)\,dx \qquad ;\forall y \in \mathbb{R}$$

بشكلٍ مماثل لما سبق، وتحت الفرض أنّنا رمزنا للنهاية $\lim_{h\to 0} P\left(Y < y \mid x \le X < x+h\right)$ بالرمز $F_{Y\mid X}(y \mid x)$، فإنَّه سيكون لدينا:

$$f_{Y\mid X}(y \mid x) = \frac{f_{\mathbb{X}_2}(x , y)}{f_X(x)} \quad ;\forall y \in \mathbb{R} \qquad [5,71\text{-b}]$$

ومن العلاقتين الأخيرتين ينتج لدينا ما يلي:

$$\int_{-\infty}^{+\infty} f_Y(y) F_{X\mid Y}(x \mid y)\, dy = \int_{-\infty}^{x}\int_{-\infty}^{+\infty} f_{\mathbb{X}_2}(t , y)\, dy\, dt = F_X(x) \qquad [5,72\text{-a}]$$

وكذلك:

$$\int_{-\infty}^{+\infty} f_X(x) F_{Y\mid X}(y \mid x)\, dx = \int_{-\infty}^{y}\int_{-\infty}^{+\infty} f_{\mathbb{X}_2}(x ,u)\, dx\, du = F_Y(y) \qquad [5,72\text{-b}]$$

إنَّ الصيغتين الأخيرتين [5,72-a] و [5,72-b] يُنظر إليهما على أنَّهما تعميم لصيغة دستور الاحتمال التام الواردة في الفصل الثاني، ولكن من أجل متغيِّرات عشوائية مستمرَّة.

**تعميم:**

من أجل الحالة $3 \le n$ تُعمَّم بشكل مماثل تماماً لما سبق، فعلى سبيل المثال لو أخذنا $\mathbb{X}_3 = (X , Y , Z)$ متَّجهاً عشوائياً فإنَّه سيكون لدينا:

$$\begin{aligned} P\left(X < x \mid Y = y ,\ Z = z\right) &= \lim_{\substack{h\to 0\\ k\to 0}} P\left(X < x \mid y < Y \le y+h \ , z \le Z < z+k\right) \\ &= \frac{1}{f_{Y,Z}(y , z)} \int_{-\infty}^{x} f_{\mathbb{X}_3}(t , y , z)\, dt \end{aligned} \qquad [5,73]$$

وهنا يجب الأخذ بالحسبان ضرورة استمرار الدالّة $f_{\mathbb{X}_3}(x , y , z)$ وكذلك $f_{Y,Z}(y , z)$ وأنَّ يكون $0 < f_{Y,Z}(y , z)$ من أجل كل $(y , z) \in \mathbb{R}^2$. علماً أنَّ $f_{Y,Z}(y , z)$ هي دالَّة الكثافة الهامشيّة لـ $Y$ و $Z$.

### (١, ٣, ٦, ٥) دوال الكثافة الشرطية لمتَّجه عشوائي مستمرّ

ليكن $\mathbb{X}_3 = (X , Y , Z)$ متَّجهاً عشوائياً مستمراً بدالَّة كثافة احتمالية $f_{\mathbb{X}_3}$، ولنأخذ $x , y , z \in \mathbb{R}$ قيماً كيفية، فعندئذٍ:

١- بفرض أنَّ الحادث الآتي:

$$A := \left\{\omega \in \Omega \ ; Y(\omega) = y ,\ Z(\omega) = z\right\} \qquad \textbf{(1)}$$

مُحقَّق الوقوع باحتمال موجب تماماً (أي إنَّ $0 < P(A)$)، فإنَّ دالَّة الكثافة الشرطية لـ $X$ تحت الفرض بتحقُّق وقوع الحادث $A$ تُعرَّف بالعلاقة الآتية:

$$f_{X\mid Y,Z}(x \mid y ,z) = \frac{f_{\mathbb{X}_3}(x , y , z)}{f_{Y,Z}(y , z)} \qquad [5,74\text{-a}]$$

٢- بفرض أنَّ كلاً من الحادثين الآتيين:

$$B := \{\omega \in \Omega \mid Y(\omega) = y\} \qquad (2)$$

$$C := \{\omega \in \Omega \mid Z(\omega) \in \mathbb{R}\} \qquad (3)$$

مُحقَّق الوقوع باحتمال موجب تماماً (أي إنَّ $0 < P(B)$ و $0 < P(C)$)، فإنَّ دالَّة الكثافة الشرطية لـ $X$ وتحت الفرض بتحقُّق وقوع الحادثين $B$ و $C$ تُعرَّف بالعلاقة الآتية:

$$f_{X|Y,Z}(x \mid y, \infty) = \frac{f_{X,Y}(x, y)}{f_Y(y)} \qquad [5,74\text{-b}]$$

٣- بفرض أنَّ الحادث الآتي:

$$D := \{\omega \in \Omega \mid Z(\omega) = z\} \qquad (4)$$

مُحقَّق الوقوع باحتمال موجب تماماً (أي إنَّ $0 < P(D)$)، فإنَّ دالَّة الكثافة الشرطية لـ $X$ و $Y$ وتحت الفرض بتحقُّق وقوع الحادث $D$ تُعرَّف بالعلاقة الآتية:

$$f_{X,Y|Z}(x, y \mid z) = \frac{f_{\mathbb{X}_3}(x, y, z)}{f_Z(z)} \qquad [5,75\text{-a}]$$

### (٢ , ٣ , ٦ , ٥) دوال التوزيع الشرطية لمتَّجه عشوائي مستمر

ليكن $\mathbb{X}_3 = (X, Y, Z)$ متَّجهاً عشوائياً مستمرّاً بدالَّة كثافة احتمالية $f_{\mathbb{X}_3}$، ولنأخذ $x, y, z \in \mathbb{R}$ قيماً كيفية، فعندئذ:

١- تحت الفرض بتحقُّق وقوع الحادث $A$ المُعطى بالعلاقة (1) وكان $0 < P(A)$ فإنَّ دالَّة التوزيع الشرطية لـ $X$ وتحت الفرض بتحقُّق وقوع الحادث $A$ تُعطى من أجل كل $x \in \mathbb{R}$ بوساطة العلاقة الآتية:

$$F_{X|Y,Z}(x \mid y, z) = P(X < x \mid A)$$

وتُكتب عادة على النحو الآتي:

$$F_{X|Y,Z}(x \mid y, z) = \int_{-\infty}^{x} f_{X|Y,Z}(t \mid y, z)\, dt$$

٢- تحت الفرض بتحقُّق وقوع الحادثين $B$ و $C$ المُعطيين بالعلاقتين **(2)** و**(3)** وكان $0 < P(B)$ وكذلك $0 < P(C)$ فإنَّ دالَّة التوزيع الشرطية لـ $X$ وتحت الفرض بتحقُّق وقوع الحادثين $B$ و $C$ تُعطى من أجل كل $x \in \mathbb{R}$ بوساطة العلاقة الآتية:

$$F_{X|Y,Z}(x \mid y, z) = P(X < x \mid B, C)$$

وتُكتب عادة على النحو الآتي:

$$F_{X|Y,Z}(x \mid y, \infty) = \int_{-\infty}^{x} f_{X|Y,Z}(t \mid y, \infty)\, dt$$

٣- تحت الفرض بتحقُّق وقوع الحادث $D$ المُعطى بالعلاقة **(4)** وكان $0 < P(D)$ فإنَّ دالَّة التوزيع الشرطية لـ $X$ و $Y$ وتحت الفرض بتحقُّق وقوع الحادث $D$ تُعطى من أجل كل $x, y \in \mathbb{R}$ بوساطة العلاقة الآتية:

$$F_{X,Y|Z}(x, y \mid z) = P(X < x, Y < y \mid D)$$

وتُكتب عادة على النحو الآتي:

$$F_{X,Y|Z}(x, y \mid z) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y|Z}(t, u \mid z)\, dt\, du$$

**(٣ , ٣ , ٦ , ٥) أمثلة**

١- بالعودة إلى المثال / ٢/ من (٤ , ٢ , ٥ , ٥) سنبحث في تعيين دالَّة التوزيع الهامشيّة للمتغيِّر العشوائي $X$ في توزيع المتَّجه العشوائي $\mathbb{X}_2 = (X , Y)$ ، وكذلك تعيين دالَّة التوزيع الشرطية للمتغيِّر العشوائي $X$ عندما يكون الحادث $\{Y = y\}$ مُحقَّقاً باحتمال موجب تماماً.

الحـل: لنقم أولاً بحساب دالَّة الكثافة الهامشيّة للمتغيِّر العشوائي $X$، فنجد لها العرض الآتي من أجل كل $x \in \mathbb{R}$ :

$$f_X(x) = \int_{-\infty}^{\infty} f_{\mathbb{X}_2}(x , y)\, dy = \begin{cases} 0 & for \;\; x \le 0 \;\; or \;\; x \ge 1 \\ 2x & for \;\; 0 < x < 1 \end{cases}$$

ومن ثمَّ يكون لدالَّة التوزيع الهامشيّة لـ $X$ في توزيـع $\mathbb{X}_2$ العرض الآتي:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt = x^2 \cdot \mathbf{I}_{(0,1)}(x) + \mathbf{I}_{[1,\infty)}(x)$$

الآن لتعيين دالَّة التوزيع الشرطية لـ $X$ عندما يكون الحادث $\{Y = y\}$ مُحقَّقاً باحتمال موجب تماماً سنقوم أولاً بتعيين دالَّة الكثافة الشـرطية لـ $X$ عندما يكون الحـادث $\{Y = y\}$ مُحقَّقاً باحتمال موجب تماماً، فنجد من أجـل أي $y \in (0 , 1)$ ما يلي:

$$f_{X|Y}(x \mid y) = \begin{cases} 2x & for \;\; 0 < x < 1 \\ 0 & otherwise \end{cases}$$

وأمّا من أجل قيم $y$ التي لا تنتمي إلى الفترة $(0 , 1)$ سيكون لدينا $f_Y(y) = 0$ ، ومن ثمَّ يصبح لدينا في هذه الحالة:

$$f_{X|Y}(x , y) = 0$$

وهكذا يكون للدالَّة $F_{X|Y}(x \mid y)$ العرض الآتي:

$$F_{X|Y}(x \mid y) = \int_{-\infty}^{x} f_{X|Y}(t \mid y)\, dt = x^2 \cdot \mathbf{I}_{(0,1)}(x) + \mathbf{I}_{[1,\infty)}(x) \qquad ; \forall y \in (0 , 1)$$

علماً أنَّه من أجل قيم $y$ التي لا تنتمي إلى الفترة $(0 , 1)$ سيكون لدينا $F_{X|Y}(x , y) = 0$ ، وبذلك يكون قد تمَّ المطلوب.

٢- ليكن $\mathbb{X}_2 = (X , Y)$ متَّجهاً عشوائياً مستمرّاً بدالَّة كثافة احتمالية

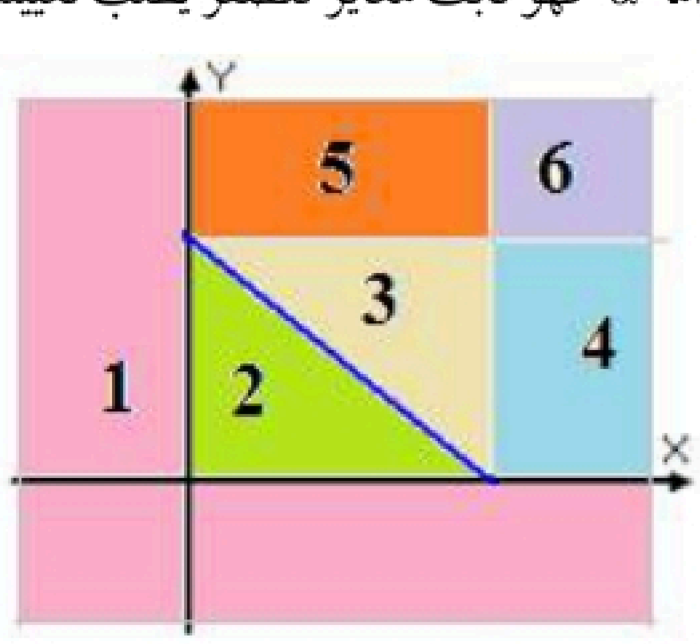
$$f_{\mathbb{X}_2}(x , y) = \begin{cases} \alpha x y & for \;\; (x , y) \in D \\ 0 & otherwise \end{cases}$$

علماً أنَّ $D$ هي الساحة المُحدَّدة من خلال العلاقات $x = 0$ و $y = 0$ و $x + y = 1$ ، وأمّا $\alpha$ فهو ثابت مغاير للصفر يُطلب تعيينه، ولنقم بتعيين $f_{X|Y}(x \mid y)$ ومن ثمَّ $F_X(x)$ ، وأخيراً تعيين $F_{X|Y}(x , y)$ .

الحـل: من أجل ذلك لنقم بتجزئة الساحة $\mathbb{R}^2$ إلى ساحات جزئية كما هو موضَّح في الشكل الجانبي (٥ , ٢٩)، حيث نلاحظ أنَّ الساحة $D$ ممثَّلة بالقطاع (2)، ومن ثمَّ لحساب الثابت $\alpha$ لدينا:

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{\mathbb{X}_2}(x , y)\, dx\, dy = \alpha \int_0^1 dy \int_0^{1-y} x\, y\, dx = \frac{\alpha}{24}$$

ومنه ينتج أنَّ $\alpha = 24$ ، علماً أنَّ التكاملات التي قيمها معدومة لم تُكتب في الصيغة الأخيرة.



الشكل (٥ , ٢٩)

الآن لتعيين دالَّة توزيع $\mathbb{X}_2$ نجد ما يلي:

أ- من أجل كل $(x\,,\,y)$ من الساحة رقم **(1)** لدينا:

$$F_{\mathbb{X}_2}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f_{\mathbb{X}_2}(t,u)\;dt\,du\;=0$$

ب- من أجل كل $(x\,,\,y)$ من الساحة رقم **(2)** لدينا:

$$F_{\mathbb{X}_2}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f_{\mathbb{X}_2}(t,u)\;dt\,du\;=6\,x^2y^2$$

ج- من أجل كل $(x\,,\,y)$ من الساحة رقم **(3)** لدينا:

$$F_{\mathbb{X}_2}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f_{\mathbb{X}_2}(t,u)\;dt\,du=1-\left[\int_{x}^{1}\int_{0}^{1-x} f_{\mathbb{X}_2}(t,u)\;dt\,du+\int_{y}^{1}\int_{0}^{1-y} f_{\mathbb{X}_2}(t,u)\;dt\,du\right]$$

وبالتعويض عن $f_{\mathbb{X}_2}(t,u)$ بما يساويها ومن ثمَّ المكاملة نحصل على الصيغة الآتية لدالَّة التوزيع $F_{\mathbb{X}_2}(x,y)$ على هذه الساحة:

$$F_{\mathbb{X}_2}(x,y)=1-24\left[\int_{x}^{1} t\,dt\int_{0}^{1-x} u\,du+\int_{y}^{1} u\,du\int_{0}^{1-y} t\,dt\right]=1-6\left[(1-x^2)\cdot(1-x)^2+(1-y^2)\cdot(1-y)^2\right]$$

د- من أجل كل $(x\,,\,y)$ من الساحة رقم **(4)** لدينا:

$$F_{\mathbb{X}_2}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f_{\mathbb{X}_2}(t,u)\;dt\,du=24\int_{0}^{y} u\;du\int_{0}^{1-y} t\;dt=6y^2-8y^3+3y^4$$

هـ- من أجل كل $(x\,,\,y)$ من الساحة رقم **(5)** لدينا:

$$F_{\mathbb{X}_2}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f_{\mathbb{X}_2}(t,u)\;dt\,du=24\int_{0}^{x} t\;dt\int_{0}^{1-x} u\;dt=6x^2-8x^3+3x^4$$

و- من أجل كل $(x\,,\,y)$ من الساحة رقم **(6)** لدينا:

$$F_{\mathbb{X}_2}(x,y)=\int_{-\infty}^{x}\int_{-\infty}^{y} f_{\mathbb{X}_2}(t,u)\;dt\,du=24\int_{0}^{1} u\;du\int_{0}^{1-y} t\;dt=1$$

ومن ثمَّ ينتج لدينا أنَّ لدالَّة التوزيع $F_{\mathbb{X}_2}(x,y)$ العرض الآتي:

$$F_{\mathbb{X}_2}(x,y)=\begin{cases} 0 & for\ x\le 0\ or\ y\le 0 \\ 6\,x^2y^2 & for\ 0<x\le 1\ and\ 0<y\le 1\ with\ x+y\le 1 \\ 1-6\left[(1-x^2)\cdot(1-x)^2+(1-y^2)\cdot(1-y)^2\right] & \\ & for\ 0<x\le 1\ and\ \ 0<y\le 1\ with\ x+y>1 \\ 6y^2-8y^3+3y^4 & for\ x>1\ and\ \ 0<y\le 1 \\ 6x^2-8x^3+3x^4 & for\ \ 0<x\le 1\ and\ \ y>1 \\ 1 & for\ x>1\ and\ \ y>1 \end{cases}$$

و الآن لتعيين دالَّة التوزيع الهامشيّة لـ $X$ يجب علينا تعيين دالَّة كثافته الهامشيّة $f_X(x)$ أولاً حيث لدينا من أجل كل $x \in \mathbb{R}$ ما يلي:

$$f_X(x) = \int_{-\infty}^{\infty} f_{X_2}(x, y)\, dy = \begin{cases} 12x(1-x)^2 & for\ 0 < x \le 1 \\ 0 & otherwise \end{cases}$$

ومن ثمَّ يكون لدالَّة التوزيع الهامشيّة لـ $X$ العرض الآتي:

$$F_X(x) = \int_{-\infty}^{x} f_X(t)\, dt = \begin{cases} 0 & for\ x \le 0 \\ 6x^2 - 8x^3 + 3x^4 & for\ 0 < x \le 1 \\ 1 & for\ x > 1 \end{cases}$$

أخيراً لتعيين دالَّة التوزيع الشرطية لـ $X$ عندما يكون الحادث $\{Y = y\}$ مُحَقَّقاً باحتمال موجب تماماً علينـا أولاً تعيين الكثافة الشرطية لـ $X$، فنجد ما يلي:

أ- من أجل $0 < y \le 1$ لدينا:

$$f_{X|Y}(x \mid y) = \frac{f_{X_2}(x, y)}{f_Y(y)} = \begin{cases} \dfrac{2x}{(1-y)^2} & for\ 0 < x \le 1\ and\ x + y \le 1 \\ 0 & otherwise \end{cases}$$

وأمَّا من أجل قيم $y$ التي لا تنتمي إلى الفترة $]0, 1]$ سيكون لدينا $f_{X|Y}(x \mid y) = 0$، ومن ثمَّ ينتج أنَّ لدالَّة التوزيع الشرطية لـ $X$ العرض الآتي من أجل كل $y \in ]0, 1]$:

$$F_{X|Y}(x \mid y) = \int_{-\infty}^{x} f_{X|Y}(t \mid y)\, dt = \begin{cases} 0 & for\ x \le 0 \\ \dfrac{x^2}{(1-y)^2} & for\ 0 < x \le 1-y \\ 1 & for\ x > 1-y \end{cases}$$

وبذلك يكون الحل قد اكتمل.

## هذا ما تيسَّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل

# تمارين الفصل الخامس

١- ليكن $[\Omega,\mathscr{A},P]$ الفضاء الاحتمالي لتجربة إلقاء حجري نرد متمايزين ومتوازنين لمّرة واحدة، وليكن $X$ تطبيقاً حقيقياً معرَّفاً على $\Omega$ كما يلي:

$$X(\omega)=X((i\ ,\ j)):=\begin{cases}+1 & for\ \ i+j\geq 10\\ 0 & for\ \ 1\leq i<3\ \ and\ \ j\in\{1,\dots 6\}\\ -1 & for\ \ i=3\ \ and\ \ j\leq 3\\ -5 & otherwise\end{cases}$$

والمطلوب ما يلي:

أ - أثبت أنَّ $X$ هو متغيِّر عشوائي فوق $[\Omega,\mathscr{A},P]$ .

ب- عيِّن دالّة توزيع $X$ وارسمها.

جـ- احسب الاحتمال $P(-3<X\leq 0)$ .

د - هل يبقى $X$ متغيِّرا عشوائيًا فوق $[\Omega,\mathscr{A},P]$ إذا كان حجري النرد غير متمايزين؟

٢- لدى عمليّة تصنيع منتج من نوع مُحدَّد نقوم بسحب عشوائي لمجموعات كل منها مكوَّنة من ثلاثة عناصر منتجة؛ وذلك بغية تفتيشها والتحقّق من أنَّها تحقّق المواصفات المطلوبة، فإذا كنّا ننظر إلى حصولنا على مكوِّن مَعيب أنَّه نجاح Success، وبفرض أنَّنا سحبنا 8 مجموعات وكانت نتائجها كما في الجدول الآتي (علماً أنَّ الحرف F يعني الفشل Failure):

| نتائج السحب | SSS | SFS | SSF | FSS | SFF | FSF | FFS | FFF |
|---|---|---|---|---|---|---|---|---|

فلو كان $X$ تطبيقاً راصداً لعدد مرّات النجاح في السحوبات جميعاً، فعندئذ:

أ - أثبت أنَّ التطبيق $X$ هو متغيِّر عشوائي فوق الفضاء الاحتمالي لإنتاج المصنع (يطلب تعيينه).

ب- مثِّل هذا المتغيِّر العشوائي بيانياً.

ج- عيِّن دالّة توزيعه الاحتمالية.

٣- ليكن $X$ متغيِّراً عشـوائياً مستمرَّاً بدالّة كثافة احتمالية $f_X(x)$ مُعطاة من خلال المعادلة الآتية:

$$f_X(x)=\alpha x^2\,\mathbf{I}_{[0,1]}(x)\qquad ;\forall x\in\mathbb{R}$$

والمـطلوب ما يلي:

أ- عيِّن قيمة الثابت $\alpha$، ومن ثمَّ احسب الاحتمال $P(X\leq 0.5)$ .

ب- عيِّن دالّة توزيع $X$ .

٤- لتكن $X$، $Y$، $X_1$، $X_2$،....، $X_n$،.... متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ ، فعندئذ أثبت (متن المبرهنة (٢، ٤، ١، ٥)) أنَّ الدالّة $Z$ في الصيغ الآتية هي متغيِّر عشوائي فوق $[\Omega,\mathscr{A},P]$ أيضاً:

a) $Z=|X|$ & b) $Z=aX+b\ \ ;\ a,b\in\mathbb{R}\ \ and\ \ a\neq 0$

c) $Z=X\pm Y$ & d) $Z=X\cdot Y$

e) $Z=\dfrac{X}{Y}\ \ ;Y(\omega)\neq 0\ \ for\ \ all\ \ \omega\in\Omega$ & f) $Z=\max\{X_i,X_j\}\ \ ;i,j\in\mathbb{N},i\neq j$

g) $Z = \mathbf{min}\left\{X_i , X_j\right\} \quad ; i , j \in \mathbb{N}, i \neq j$ & h) $Z = \mathbf{sup}\{X_n \; ; n \in \mathbb{N}\}$

i) $Z = \mathbf{inf}\{X_n \; ; n \in \mathbb{N}\}$ & j) $Z = \overline{\mathbf{lim}}_{n\to\infty} X_n = \mathbf{lim}_{n\to\infty}\, \mathbf{sup}_{m\geq n} X_m$

k) $Z = \underline{\mathbf{lim}}_{n\to\infty} X_n = \mathbf{lim}_{n\to\infty}\, \mathbf{inf}_{m\geq n} X_m$ & l) $Z = \mathbf{lim}_{n\to\infty} X_n$

٥- لتكن لدينا الدالَّة الآتية:

$$f(x) = c\binom{3}{x}\binom{4}{4-x} \quad ; x = 0, 1, 2, 3$$

والمطلوب تعيين قيمة الثابت $c$ بحيث تصبح هذه الدالَّة هي دالَّة الاحتمالات الكتلية لمتغيِّر عشوائي $X$.

٦- ليكن $X$ متغيِّراً عشـوائياً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ بدالَّة توزيع احتمالية $F_X$، والمطلوب تعيين دالَّة توزيع المتغيِّر العشوائي $Y$ في كلٍّ من الحالات الآتية:

a) $Y := aX + b \quad ; a, b \in \mathbb{R}, a \neq 0$ & b) $Y := \left| X \right|$

**c)** $Y := X^2$ & **d)** $Z := \dfrac{1}{X} \quad ; X(\omega) \neq 0 , \forall \omega \in \Omega$

٧- ليكن $X$ متغيِّراً عشـوائياً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ بدالَّة توزيع احتمالية $F_X$، والمطلوب إثبات أنَّ $F_X$ هي متغيِّر عشوائي أيضاً، ومن ثمَّ تعيين دالَّة توزيعها، أي تعيين صيغة الدالَّة $F_{F_X}$.

٨- ليكن $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً مُعطى على النحو الآتي:

$$\Omega = [0 , \infty) \times [0 , \infty) \qquad \& \qquad \mathscr{A} = \Re^2 \cap \{[0 , \infty) \times [0 , \infty)\}$$

$$\text{and} \quad P(A) := \int_A \mathbf{e}^{-(x_1 + x_2)}\, dx_1\, dx_2 \quad ; \forall A \in \mathscr{A}$$

وليكن $X$ تطبيقاً من $\Omega$ في $\mathbb{R}$ مُعرَّفاً بالعلاقة $X((x_1, x_2)) := x_1 + x_2$، والمطلوب إثبات أنَّ $X$ هو متغيِّر عشوائي فوق $[\Omega, \mathscr{A}, P]$.

٩- ليكن $[\Omega, \mathscr{A}, P]$ فضاء احتمالي، علماً أنَّ:

$$\Omega = \mathbb{N}_6 \qquad \& \qquad \mathscr{A} = 2^{\Omega} \qquad \& \qquad P(\{\omega\}) = \frac{1}{6}$$

ولنأخذ $\mathbb{X}_2 = (X , Y)$ متَّجهاً عشوائيًّا متقطِّعًا فوق $[\Omega, \mathscr{A}, P]$ بحيث إنَّه من أجل كل $\omega \in \Omega$ لدينا:

$$X(\omega) = \omega \qquad \& \qquad Y(\omega) = 5\omega + 1$$

والمطلوب تعيين دالَّة التوزيع للمتَّجه العشوائي $\mathbb{X}_2$.

١٠- ليكن $[\Omega, \mathscr{A}, P]$ فضاء احتمالي، علماً أنَّ:

$$\Omega = \mathbb{N}_6 \qquad \& \qquad \mathscr{A} = 2^{\Omega} \qquad \& \qquad P(\{\omega\}) = \frac{1}{6}$$

ولنأخذ $\mathbb{X}_3 = (X , Y , Z)$ متَّجهاً عشوائيًّا متقطِّعًا فوق $[\Omega, \mathscr{A}, P]$ وبحيث إنه من أجل كل $\omega \in \Omega$ لدينا:

$$X(\omega) = \omega \qquad \& \qquad Y(\omega) = [X(\omega)]^2 \qquad \& \qquad Z(\omega) = [X(\omega)]^3(\omega) \qquad ; \forall \omega \in \Omega$$

والمطلوب تعيين دالَّة التوزيع للمتَّجه العشوائي $\mathbb{X}_3$.

١١- لدينا صندوق يحوي 24 كرة متماثلة تماماً ومرقّمة بالأرقام 1 وحتى 24. تُخلط الكرات جيداً ومن ثمّ تُسحب كرة من الصندوق دون إعادتها، وليكن $\mathbb{X}_3 = (X, Y, Z)$ متّجهاً عشوائياً فوق الفضاء الاحتمالي لهذه التجربة العشوائيّة بحيث إنّ:

$$X(\omega) = \begin{cases} 0 & for\ \omega\ even \\ 1 & for\ \omega\ odd \end{cases}$$

$$Y(\omega) = \begin{cases} 0 & for\ \omega\ divisible\ on\ 3 \\ 1 & otherwise \end{cases}$$

$$Z(\omega) = \begin{cases} 0 & for\ \omega \le 12 \\ 1 & for\ \omega > 12 \end{cases}$$

والمطلوب ما يلي:

أ- تعيين دالّة توزيع هذا المتّجه العشوائي.

ب- تعيين التوزيعات الهامشيّة لمركبات هذا المتّجه العشوائي.

ج- تعيين التوزيعات الشرطية $p_{i\,|\,j\,k}$ و $p_{i\,j\,|\,k}$ من أجل $i = 1$ و $j = 2$ وأخيراً $k = 1$.

١٢- ليكن $\mathbb{X}_2 = (X, Y)$ متّجهاً عشوائياً مستمرّاً بدالّة كثافة احتمالية مُعطاة من خلال العلاقة الآتية:

$$f_{\mathbb{X}_2}(x, y) = \begin{cases} \alpha(x \cdot y) & for\ \ (x, y) \in B \\ 0 & otherwise \end{cases}$$

علماً أنّ $\alpha \neq 0$ ثابت عـددي، والمطلوب ما يلي:

أ- تعيين قيمة الثابت $\alpha$.

ب- تعيين دالّة توزيع المتّجه العشوائي $\mathbb{X}_2$ في كل من الحالتين الآتيتين:

- إذا كانت المجموعة $B$ هي المجال $(0, 1) \times (0, 1)$.

- إذا كانت المجموعة $B$ هي الساحة المحدّدة بالعلاقات $x = 0$ و $y = 0$ و $x + y = 1$

١٣- ليكن $\mathbb{X}_2 = (X, Y)$ متّجهاً عشوائياً مختلطاً فيه $X$ متغيّر عشوائي مستمر بدالّة كثافة احتمالية:

$$f_X(x) = \lambda\, \mathbf{e}^{-\lambda x} \cdot \mathbf{I}_{[0, \infty)}(x) \qquad ; \forall x \in \mathbb{R}$$

وأمّا $Y$ فهو متغيّر عشوائي متقطّع بدالّة احتمالية كتلية مُعطاة كما يلي:

$$P(Y = y_j) = p_j = \frac{\alpha^j}{j\,!}\, \mathbf{e}^{-\alpha} \qquad ; \forall j \in \mathbb{N}^o$$

فإذا علمنا أنّ $0 < \lambda$ و $0 < \alpha$ قيم حقيقية ثابتة، وأنّ دالّة الكثافة الاحتمالية للمتّجه العشوائي $\mathbb{X}_2$ مُعطاة من خلال العلاقة الآتية:

$$f_{\mathbb{X}_2}(x, y_j) = \lambda\, \mathbf{e}^{-\lambda x} \cdot \frac{\alpha^j}{j\,!}\, \mathbf{e}^{-\alpha} \cdot \mathbf{I}_{[0, \infty)}(x)$$

فعندئذ عيّن دالّة توزيع المتّجه العشوائي $\mathbb{X}_2$.

**١٤-** ليكن $\mathbb{X}_2 = (X\ ,\ Y\ )$ متَّجهاً عشوائياً مستمرّاً بدالَّة كثافة احتمالية مُعطاة بالعلاقة الآتية:

$$f_{\mathbb{X}_2}(x\,,y\,) = \begin{cases} \alpha(x+y) & for\ (x\,,y\,) \in A \\ 0 & otherwise \end{cases}$$

علماً أنَّ $A$ هي الساحة $[0\,,1]\times[0\,,x\,]$ ، وأمّا $\alpha$ فهو ثابت عددي. عندئذ المطلوب ما يلي:

أ- عيِّن قيمة الثابت $\alpha$ .

ب- عيِّن دالَّة التوزيع لمتَّجه العشوائي $\mathbb{X}_2$ .

ج- عيِّن دوال التوزيع الهامشيّة لمركبات هذا المتَّجه العشوائي.

د- عيِّن دوال التوزيع الشرطية لمركبات هذا المتَّجه العشوائي.

**١٥-** ليكن $\mathbb{X}_2 = (X\ ,\ Y\ )$ متَّجهاً عشوائياً مستمرّاً بدالَّة كثافة احتمالية مُعطاة بالعلاقة الآتية:

$$f_{\mathbb{X}_2}(x\,,y\,) = \begin{cases} \mathbf{e}^{\alpha(x+y)} & for\ (x\,,y\,) \in A \\ 0 & otherwise \end{cases}$$

علماً أنَّ المجموعة $A$ هي المجال $(0\,,1)\times(0\,,1)$ ، وأمّا $\alpha$ فهو ثابت عددي، والمطلوب مايلي:

أ- عيِّن قيمة الثابت $\alpha$ .

ب- عيِّن دالَّة التوزيع الاحتمالية للمتَّجه العشوائي $\mathbb{X}_2$ .

ج- عيِّن دوال التوزيع الهامشيّة لمركبات هذا المتَّجه العشوائي.

د- عيِّن دوال التوزيع الشرطية لمركِّبات هذا المتَّجه العشوائي.

# الفصل السادس

# توزيعات احتمالية شهيرة

## FAMOUS PROBABILITY DISTRIBUTIONS

سنخصِّص هذا الفصل لتقديم بعض التوزيعات الاحتمالية الشهيرة التي لها دور مهمّ في الكثير من الدراسات الاحتمالية والإحصائية على حدٍّ سواء وسنبدأها بالتوزيعات المتقطّعة.

## (٦,١) توزيعات احتمالية متقطعة

**Discrete Probability Distributions**

فيما يلي نقدِّم بعض التوزيعات الاحتمالية المتقطِّعة الشهيرة التي نستهلّها بأبسط نوع منها ألا وهو التوزيع الآتي:

### (٦ , ١ , ١) التوزيع الوحيد النقطة One Point Distribution

إنَّ سبب تسمية هذا التوزيع بهذا الاسم يعود إلى أنَّ مجموعة قيم المتغيِّر العشوائي $X$ الواصف لهذا التوزيع تحتوي على قيمة وحيدة فقط (ويُعرف باسم التوزيع المُضْمَحل **Degenerate Distribution** أيضاً).

### (٦ , ١ , ١ , ١) تعريفه

ليكن $X$ متغيِّراً عشوائياً متقطِّعاً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، فعندئذ يُقال عن $X$ إنَّه خاضع للتوزيع **وحيد النقطة** إذا كان $X$ ثابتاً باحتمال يساوي الواحد، أي إنه إذا وجد $c \in \mathbb{R}$ بحيث تكون العلاقة الآتية مُحقَّقة:

$$P(X = c) = 1 \qquad [6,1]$$

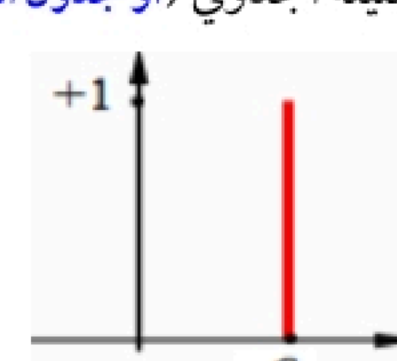
إذاً، فمن الملاحظ أنَّ مجموعة قيم هذا المتغيِّر العشوائي $\mathbf{X} = \{c\}$ (قدِّم كمثال في الفصل السابق) تحتوي على قيمة وحيدة $c$ فقط، وهذه القيمة تُدعى **مَعْلمة التوزيع** Parameter of Distribution. كذلك يُعرف هذا المتغيِّر العشوائي باسم **المتغيِّر العشوائي الثابت** ويقابل نموذج الدالَّة الثابتة في الرياضيات.

### (٦ , ١ , ١ , ٢) ملاحظات

١ – يستخدم هذا النوع من المتغيِّرات العشوائية في المسائل التي تقترن (أو ترتبط) فيها كل نتائج التجرب العشوائية بقيمة ثابتة واحدة فقط، وتمثيله الجدولي (أو جدول التوزيع) والبياني موضَّح كما يلي:

الجدول (٦ , ١)

| $i$ | 1 | sum |
|---|---|---|
| قيم $X$ $x_i$ | $c$ | |
| $p_i = P(X = x_i)$ | 1 | 1 |



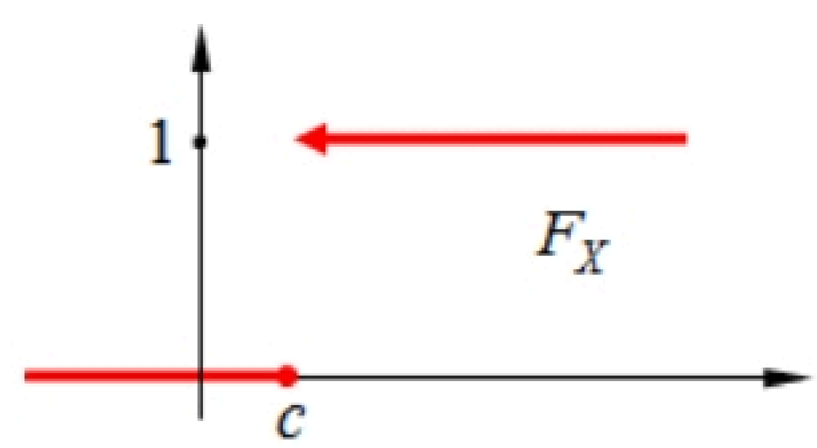

الشكل (٦ , ١). التمثيل البياني لـ $X$.

٢- إنَّ دالَّة توزيع هذا المتغيِّر العشوائي نجدها باستخدام تعريف دالَّة التوزيع لمتغيِّر عشوائي متقطِّع، حيث لدينا:

$$F_X(x) = \sum_{i \in I\,;\, x_i < x} P(X = x_i) = \mathbf{I}_{(c,+\infty)}(x) \qquad ;x \ni \mathbb{R}$$

كما يمكن كتابتها وفقاً للصيغ الشرطية على النحو الآتي أيضاً:

$$F_X(x) = \begin{cases} 0 & for\ x \le c \\ 1 & for\ x > c \end{cases} \quad \textbf{or} \quad F_X(x) = \begin{cases} 0 & if\ \ x \le c \\ 1 & if\ \ x > c \end{cases} \quad \textbf{or} \quad F_X(x) = \begin{cases} 0 & for\ x \le c \\ 1 & otherwise \end{cases}$$




الشكل (٦,٢) العرض البياني لدالَّة توزيع $X$.

فنلاحظ أنَّ دالَّة توزيع هذا المتغيِّر العشوائي هي أبسط أنواع الدوال الدرجية، ولذلك يُعدّ هذا المتغيِّر العشوائي من أبسط المتغيِّرات العشوائية على الإطلاق.

٣- سوف نخصص رمزاً لكل متغير عشوائي شهير يتمّ تقديمه، فعلى سبيل المثال سنرمز للتوزيع الوحيد النقطة (التوزيع المضمحل) بالرمز $\mathbf{D}(c)$ .

٤- عندما نكتب $X \sim \Lambda(\ldots)$ نقصد بذلك أنَّ المتغيِّر العشوائي $X$ خاضع لتوزيع رمزه $\Lambda(\ldots)$، فعلى سبيل المثال عندما نكتب $X \sim \mathbf{D}(c)$ فإنَّنا نقصد بذلك أنَّ المتغيِّر العشوائي $X$ خاضع للتوزيع الوحيد النقطة في $c$ (أو بمَعْلَمَة $c$)، وهذا يعني أنّ العلاقة [6,1] مُحقَّقة من أجل المتغيِّر العشوائي $X$ .

### (٦,١,١,٣) أمثلة

١-ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً، علماً أنَّ $\Omega$ هي مجموعة كل التباديل للأعداد 1، 2 و3، أي إنَّ:

$$\Omega := \{\underbrace{123}_{\omega_1},\underbrace{132}_{\omega_2},\underbrace{213}_{\omega_3},\underbrace{231}_{\omega_4},\underbrace{312}_{\omega_5},\underbrace{321}_{\omega_6}\}$$

وأمّا $\mathscr{A}$ سنأخذه موافقاً للجبر القوي فوق $\Omega$، أي إن $\mathscr{A} = 2^\Omega$ ، وأخيراً سنأخذ الدالَّة الاحتمالية $P$ كما يلي:

$$P(A) = \sum_{i=1\,;\,\omega_i \in A}^{6} Q(\{\omega_i\}) \qquad A \in \mathscr{A} \quad with \ \sum_{i=1}^{6} Q(\{\omega_i\}) = 1$$

ولنأخذ $X$ تطبيقاً حقيقياً معرَّفاً فوق $\Omega$ بحيث يُقرن كل حادث ابتدائي $\omega_i$ بعدد يساوي مجموع الأعداد المكوِّنة لهذا الحادث الابتدائي، أي إنَّ:

$$X : \Omega \longrightarrow \mathbb{R} \quad ;\ \omega_i \mapsto X(\omega_i) = 6 \ \ ;\ \forall i \in \mathbb{N}_6$$

فنجد أنَّ هذا التطبيق هو متغيِّر عشوائي (تحقَّق من صحة ذلك) ويأخذ قيمة وحيدة هي 6، ومن ثَمَّ يكون $X \sim \mathbf{D}(6)$ (أي إن هذا المتغيِّر العشوائي خاضع للتوزيع الوحيد النقطة في $c = 6$).

٢- ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً، علماً أنَّ:

$$\Omega := \left\{ \omega_n := \frac{1}{n} \; ; n \in \mathbb{N} \right\}$$

وأمّا $\mathscr{A}$ سنأخذه موافقاً للجبر القوي فوق $\Omega$، أي إنَّ $\mathscr{A} = 2^{\Omega}$، وأخيراً سنأخذ الدالّة الاحتمالية $P$ المُعطاة كما يلي:

$$P(A) = \sum_{n=1;\, \omega_n \in A}^{\infty} Q(\{\omega_n\}) \quad ; \forall A \in \mathscr{A} \qquad with \;\; Q(\{\omega_n\}) = Q(\{1/n\}) = \frac{6}{\pi^2 n^2}$$

ولنأخذ $X$ تطبيقاً حقيقياً معرَّفاً فوق $\Omega$ بحيث يقرن كل حادث ابتدائي $\omega_n$ بالعدد $\boldsymbol{\pi}$، أي إنَّ:

$$X : \Omega \longrightarrow \mathbb{R} \quad ; \omega_n \mapsto X(\omega_n) = \boldsymbol{\pi} \quad ; \forall n \in \mathbb{N}$$

فنجد أنَّ هذا التطبيق هو متغيِّر عشوائي (تحقَّق من صحة ذلك) ويأخذ قيمة وحيدة هي $\boldsymbol{\pi}$، ومن ثمَّ نجد أنَّ $X \sim \mathbf{D}(\pi)$.

٣- ليكن لدينا فضاء احتمالي $[\Omega,\mathscr{A},P]$، علماً أنَّ:

$$\Omega = [a\,,b] \quad \& \quad \mathscr{A} = \Re \cap [a\,,b] \quad \& \quad P(A) = \frac{1}{b-a}\int_A dx \quad ; \forall A \in \mathscr{A}$$

ولنأخذ $X$ تطبيقاً حقيقياً معرَّفاً فوق $\Omega$ بحيث يقرن كل حادث ابتدائي $\omega$ من $\Omega$ بالحد الأعلى للفترة $[a\,,b]$، أي إنَّ:

$$X : \Omega \longrightarrow \mathbb{R} \quad ; \omega \mapsto X(\omega) = b$$

فنجد أنَّ هذا التطبيق هو متغيِّر عشوائي (تحقَّق من صحة ذلك) ويأخذ قيمة وحيدة هي $b$، ومن ثمَّ يكون $X \sim \mathbf{D}(b)$.

### (٦, ١, ١, ٤) ملاحظة

من الأمثلة السّابقة نلاحظ أنَّ طبيعة مجموعة نتائج التجربة العشوائية $\Omega$ لا تهمّنا بشيء منتهيةً كانت أم غير منتهية، قابلةً للعدّ أو غير قابلة للعدّ، وكذلك بالنسبة إلى احتمالات الحوادث الابتدائية، فمن غير المهمّ إن كانت متساوية الاحتمال أو غير متساوية الاحتمال. إنّما المهمّ في المتغيِّرات العشوائية المتقطِّعة هو عدد القيم التي يأخذها المتغيِّر العشوائي $X$ فقط، وأمّا احتمالات أخذ المتغيِّر العشوائي $X$ لتلك القيم فإنَّه يتعلَّق بشكل مباشر بطبيعة مجموعة نتائج التجربة العشوائية $\Omega$ وباحتمالات حوادثها الابتدائية.

## (٦, ١, ٢) التوزيع ثنائي النقطة (أو توزيع برنولي) Tow Point Distribution (or Bernoulli Distribution)

إنَّ هذا النوع من المتغيِّرات العشوائية المتقطِّعة (قدِّم أحد نماذجه كمثال في الفصل السابق) يأتي كثاني أبسط متغيِّر عشوائي متقطِّع لأنَّه يأخذ قيمتين فقط، والتعريف الآتي يوضّح لنا ذلك.

### (٦, ١, ٢, ١) تعريفه

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فعندئذ يُقال عن $X$ إنَّه خاضع للتوزيع **ثنائي النقطة** في $x_1$ و $x_2$ من $\mathbb{R}$ مع $x_1 \neq x_2$، إذا كان قانون توزيعه مُحقِّقاً للعلاقتين الآتيتين:

$$\left.\begin{aligned} P(X = x_1) &= p \\ P(X = x_2) &= 1-p \end{aligned}\right\} \qquad [6,2]$$

علماً أنَّ $0 < p < 1$ يُدعى **مَعْلمة التوزيع**.

### (٦, ١, ٢, ٢) ملاحظات

١- سنستخدم الرمز $\mathbf{B}(p)$ للدلالة على هذا التوزيع، وكذلك سنرمز بـ $\mathbf{b}(x_i\,;p)$ للدلالة على قيمة الكتلة الاحتمالية $P(X = x_i)$ لهذا المتغير العشوائي $X$ عند القيمة $x_i$، أي إنَّه من أجل $i \ni \{1, 2\}$ لدينا $\mathbf{b}(x_i\,;p) \triangleq P(X = x_i)$.

٢- في الحالة الخاصة عندما تكون $x_1 = 1$ و $x_2 = 0$ فعندئذ يُقال عن القيمة $p$ إنّها **احتمال النجاح** Success Probability، وأمّا القيمة $q := 1-p$ فتدعى **احتمال الفشل** Failure Probability، والسبب في ذلك أنّه تُعطى (**عادة**) القيمة $x_1 = 1$ عندما نحصل على النتيجة التي نصبو إليها أو نبتغيها، ولذلك يُقال إنّنا قد حصلنا على نجاح، في حين أنّ القيمة $x_2 = 0$ تُعطى (**عادة**) عندما نفشل في الحصول على النتيجة التي نصبو إليها أو نبتغيها، ولذلك يُقال إنّنا قد فشلنا في الحصول على النتيجة المرجوة. من أجل هذه الحالة الأخيرة يصبح للمتغيِّر العشوائي $X$ تسمية أخرى حيث يُقال عنه **متغيِّر عشوائي برنولي** Bernoulli Random Variable نسبة إلى الرياضياتي السويسري **برنولي** (1655-1705) Jacob Bernoulli، وهنا تجدر الإشارة إلى أنّ الرياضياتي **يعقوب برنولي** هو أول من قام بتقديم دراسة علميّة متقنة للتجارب التي تتمخض عنها إحدى نتيجتين، وسُمِّيت فيما بَعد بـ **التجارب البرنولية** Bernoulli Experiments تخليداً لذكراه، وأمّا المتغيِّرات العشوائية المرافقة لهذه التجارب فتُدعى **متغيِّرات عشوائية برنولية** كما ذُكر آنفًا. لاحظ أنّه من أجل هذا النوع من المتغيِّرات العشوائية يُنظر إلى كل نتيجة محتملة على أنّها إمّا نجاح أو فشل.

٣- يستخدم هذا النوع من المتغيِّرات العشوائية في المسائل التي تقترن فيها أيّة نتيجة من نتائج التجرب العشوائية بإحدى نتيجتين فقط، ومنها على سبيل المثال: رمي قطعة نقود لمرّة واحدة، فاعلية عقار دوائي ضدَّ مرض معيَّن، أو...

٤- التمثيل الجدولي (**جدول التوزيع**) والبياني لهذا المتغيِّر العشوائي من أجل الحالتين المذكورتين أعلاه موضَّح كما يلي:

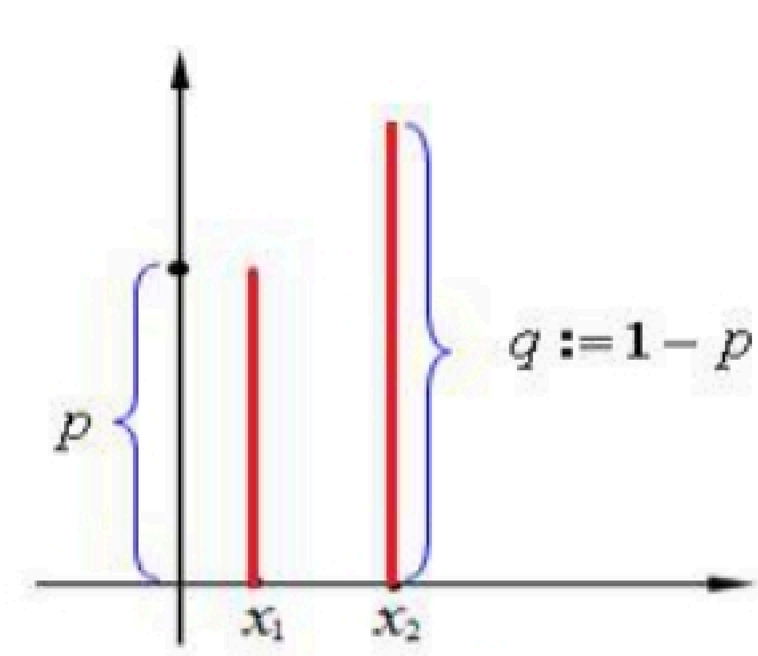
الجدول (٢, ٦. أ)

| $i$ | 1 | 2 | sum |
|---|---|---|---|
| قيم $x_i$ $X$ | $x_1$ | $x_2$ | |
| $p_i = P(X = x_i)$ | $p$ | $q := 1-p$ | 1 |

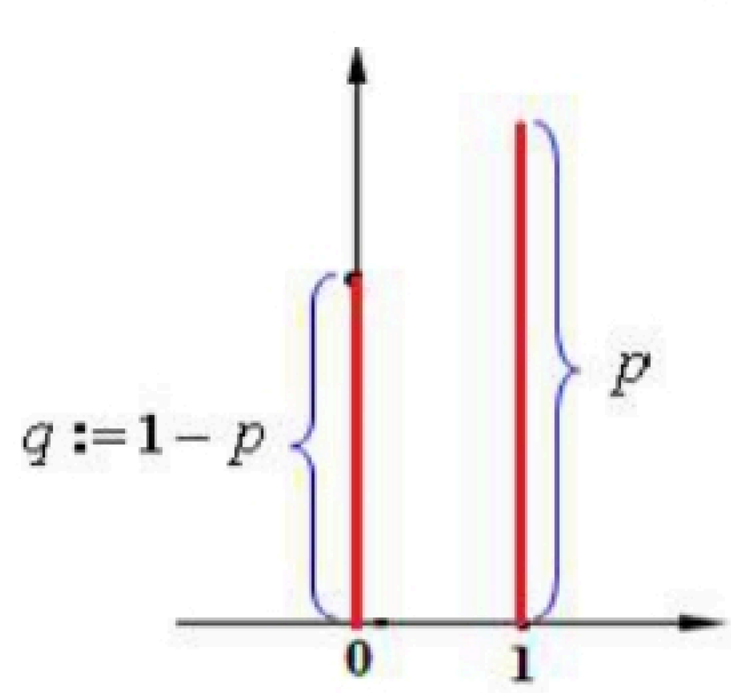
الجدول (٢, ٦. ب)

| $i$ | 1 | 2 | sum |
|---|---|---|---|
| قيم $x_i$ $X$ | 0 | 1 | |
| $p_i = P(X = x_i)$ | $q := 1-p$ | $p$ | 1 |

٥- التمثيل البياني لهذا المتغيِّر العشوائي يقدّمه لنا الشكلان الآتيان:

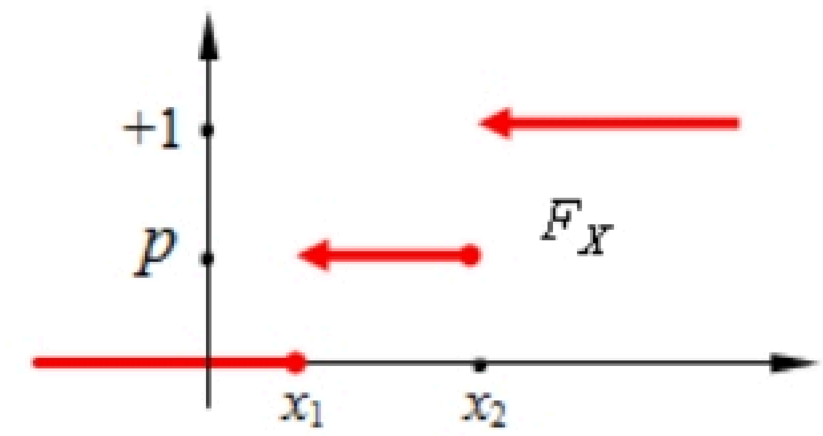



الشكل (٣, ٦. أ) من أجل $x_2 > x_1$ .




الشكل (٣, ٦. ب) من أجل الحالة البرنولية.

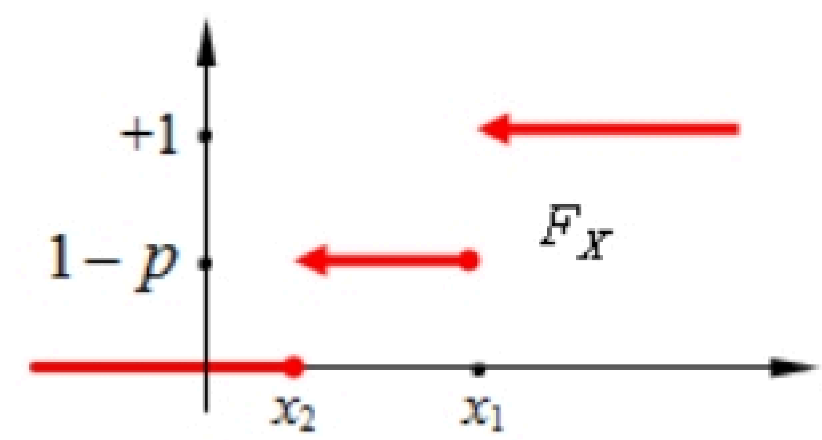
٦- إذا كان $x_2 > x_1$ فإنّه سيكون لدالّة توزيع هذا المتغيِّر العشوائي العرض والشكل الآتيين:




$$F_X(x) = \begin{cases} 0 & for\ x \le x_1 \\ p & for\ x_1 < x \le x_2 \\ p+(1-p)=1 & for\ x > x_2 \end{cases}$$

الشكل (٤ ,٦.أ) دالَّة التوزيع الاحتمالية من أجل $x_2 > x_1$ .

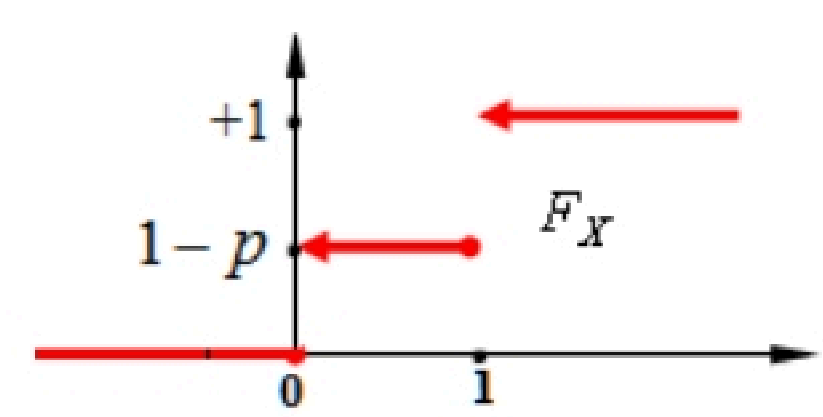
وأمّا إذا كان $x_1 > x_2$ فإنّه سيصبح لدالّة توزيع هذا المتغيِّر العشوائي العرض والشكل الآتيين:




$$F_X(x) = \begin{cases} 0 & for\ x \le x_2 \\ 1-p & for\ x_2 < x \le x_1 \\ 1 & for\ x > x_1 \end{cases}$$

الشكل (٤ ,٦.ب) دالَّة التوزيع الاحتمالية من أجل $x_1 > x_2$ .

ومن أجل الحالة البرنولية سينتج لدينا أنَّ لدالّة توزيع هذا المتغيِّر العشوائي العرض والشكل الآتيين:




$$F_X(x) = \begin{cases} 0 & for\ x \le 0 \\ 1-p & for\ 0 < x \le 1 \\ 1 & for\ x > 1 \end{cases}$$

الشكل (٤ ,٦.ج) دالَّة التوزيع الاحتمالية من أجل الحالة البرنولية.

### (٣, ٢, ١, ٦) أمثلة

١- بالعودة إلى المثال / ١/ من (٣, ١, ١, ٦) حيث لدينا:

$$\Omega := \{\underbrace{123}_{\omega_1}, \underbrace{132}_{\omega_2}, \underbrace{213}_{\omega_3}, \underbrace{231}_{\omega_4}, \underbrace{312}_{\omega_5}, \underbrace{321}_{\omega_6}\}$$

ولنأخذ $X$ تطبيقاً حقيقياً معرَّفاً فوق $\Omega$ على النحو الآتي:

$$X : \Omega \longrightarrow \mathbb{R} \ ; \omega \mapsto X(\omega) = \begin{cases} 0 & for\ \omega = 213, 312, 231 \\ 1 & for\ \omega = 123, 321, 132 \end{cases}$$

فنجد أنَّ هذا التطبيق هو متغيِّر عشوائي (تحقَّق من صحة ذلك) ويأخذ قيمتين فقط هما 0 و 1 بالاحتمال نفسه، ومن ثَمَّ يكون لهذا المتغيِّر العشوائي توزيع ثنائي النقطة في $x_1 = 1$ و $x_2 = 0$ ، وكما هو مُلاحظ فإنَّ لهذا المتغيّر العشوائي عرضًا برنوليًّا واحتمال النجاح له يساوي 0.5 ، أي إنَّ $X \sim \mathbf{B}(0.5)$ .

٢- بالعودة إلى المثال /٢/ من (٦، ١، ١، ٣) حيث لدينا $\Omega = \left\{ \omega_n = \frac{1}{n} \;\; ; n \in \mathbb{N} \right\}$، فلو أخذنا $X$ تطبيقاً حقيقياً فوق $\Omega$ معرَّفاً من خلال العلاقة الآتية:

$$X : \Omega \longrightarrow \mathbb{R} \;\; ; \omega_n \mapsto X(\omega_n) = \begin{cases} 0 & for \;\; n \;\; even \\ 1 & for \;\; n \;\; odd \end{cases}$$

لوجدنا أنَّ هذا التطبيق هو متغيِّر عشوائي (تحقَّق من صحة ذلك) ويأخذ قيمتين فقط، ومن ثمَّ يكون هذا المتغيِّر العشوائي خاضع للتوزيع ثنائي النقطة في $x_1 = 1$ و $x_2 = 0$، وبما أنَّ $P(\Omega) = 1$ فإنَّه سيكون لدينا:

$$\begin{aligned} 1 = P(\Omega) &= P\left( \left\{ \frac{1}{n} \; ; n \in \mathbb{N}, n \;\; odd \right\} + \left\{ \frac{1}{n} \; ; n \in \mathbb{N}, n \;\; even \right\} \right) \\ &= P\left( \left\{ \frac{1}{n} \; ; n \in \mathbb{N}, n \;\; odd \right\} \right) + P\left( \left\{ \frac{1}{n} \; ; n \in \mathbb{N}, n \;\; even \right\} \right) \end{aligned}$$

لأنَّ الحادثين $\left\{ \frac{1}{n} \; ; n \in \mathbb{N}, n \;\; odd \right\}$ و $\left\{ \frac{1}{n} \; ; n \in \mathbb{N}, n \;\; even \right\}$ متنافيان، وبما أنَّ $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$، فإنَّه سيكون لدينا:

$$\begin{aligned} P\left( \left\{ \frac{1}{n} \; ; n \in \mathbb{N}, n \;\; even \right\} \right) &= \sum_{n=1;\, n \;\; even}^{\infty} Q\left( \{\omega_n\} \right) = \sum_{n=1;\, n \;\; even}^{\infty} \frac{\pi^2}{6} \frac{1}{n^2} \\ &= \frac{6}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{(2n)^2} = \frac{6}{4\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{1}{4} = q \end{aligned}$$

أي إنَّ قيمة مَعْلمة هذا التوزيع $p = 0.75$، ومن ثمَّ $X \sim \mathbf{B}(0.75)$.

لاحظ هنا أنَّه لا يمكننا استنتاج قيم مثل هذه الاحتمالات باستخدام مبدأ لابلاس في الاحتمالات لأنَّ عدد الحوادث الابتدائية في أي من الحادثين غير منتهٍ فعلى سبيل المثال استخدام فكرة أنَّ الأعداد الفردية والزوجية في $\mathbb{N}$ تتناوب على التوالي فإنَّه سيكون (عددين موجبين متساويين مجموعهما الواحد) $P(X = 0) = 0.5$ و $P(X = 1) = 0.5$، ولكن الدراسة السّابقة أظهرت أنَّ هذ الاستنتاج الأخير ليس صحيحاً، بمعنى أنَّ النتائج الاحتمالية التي نحصل عليها عندما تكون $\Omega$ منتهية قد تختلف كلياً في حال أصبحت $\Omega$ غير منتهية.

٣- ليكن لدينا $[\Omega, \mathscr{A}, P]$ فضاءً احتمالياً، علماً أنَّ:

$$\Omega = [0\,,1] \quad \& \quad \mathscr{A} = \Re \bigcap [0\,,1] \quad \& \quad P(A) = \int_A dx \quad ; \forall A \in \mathscr{A}$$

ولنأخذ $X$ تطبيقاً حقيقياً معرَّفاً فوق $\Omega$ كما في العلاقة الاتية:

$$X : \Omega \longrightarrow \mathbb{R} \;\; ; \omega \mapsto X(\omega) = \begin{cases} 0 & for \;\; \omega \in \left[0\,, \frac{1}{4}\right) \\ 1 & for \;\; \omega \in \left[\frac{1}{4}, 1\right] \end{cases}$$

عندئذ نجد أنَّ:

$$\left\{ \omega \in \Omega \; ; X(\omega) < x \right\} = \begin{cases} \varnothing & for \;\; x \le 0 \\ \left[0, \frac{1}{4}\right) & for \;\; 0 < x \le +1 \\ [0,1] & for \;\; x > +1 \end{cases}$$

وبما انَّ $\varnothing$ و $\Omega = [0,1]$ هي عناصر من $\mathscr{A}$ دومًا، وكذلك لدينا $\left[0\,, \frac{1}{4}\right)$ هو عنصر من $\mathscr{A}$ أيضاً لأنَّها مجموعة بوريلية في الفترة

$[0,1]$، فإنَّ ذلك يعني إنَّ التطبيق $X$ هو متغيِّر عشوائي فوق $[\Omega,\mathcal{A},P]$ ، وبملاحظة أنَّ $X$ يأخذ قيمتين فقط، فهذا يعني أنَّ هذا المتغيِّر العشوائي خاضع للتوزيع الثنائي النقطة في $x_1=1$ و $x_2=0$، وقيمة مَعْلمته نجدها كما يلي:

$$p=P(X=1)=P\left(\left[\tfrac{1}{4},1\right]\right)=\int_{\frac{1}{4}}^{1}dx=x\Big|_{\frac{1}{4}}^{1}=1-\frac{1}{4}=\frac{3}{4}=0.75$$

ومن ثمَّ يكون $X\sim\mathbf{B}(0.75)$.

### (٦,١,٣) التوزيع المنتظم المتقطِّع Discrete Uniform Distribution

إنَّ سبب تسمية هذا التوزيع بهذا الاسم تعود إلى قيم الكتل الاحتمالية $P(X=x_i)$ لهذا المتغيِّر العشوائي حيث سنلاحظ أنَّ $X$ يأخذ جميع قيِّمه بالاحتمال نفسه (قدِّم أحد نماذجه كمثال في الفصل السابق)، أي إنَّ قيم احتمالاته منتظمة على كل القيم التي يأخذها (انظر التمثيل البياني الآتي لهذا المتغيِّر العشوائي الذي يوضِّح معنى ذلك).

#### (٦,١,٣,١) تعريفه

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$، فعندئذ يُقال عن $X$ إنَّه خاضع للتوزيع **المنتظم المتقطِّع** إذا كان يأخذ عدداً منتهياً من القيم الحقيقية $x_1$ ، $x_2$ ،... و $x_n$، وقانون توزيعه $P_X$ يُحقِّق العلاقة الآتية:

$$P(X=x_i)=\frac{1}{n}\ ;\forall\ i\in\mathbb{N}_n \qquad [6,3]$$

علماً أنَّ $n$ هو عدد صحيح موجب تماماً ويُدعى **مَعْلمة التوزيع**.

#### (٦,١,٣,٢) ملاحظة

١- سنستخدم الرمز $\mathbf{U}(n)$ للدلالة على هذا التوزيع، وسنرمز بـ $\mathbf{u}(x_i\,;n)$ للدلالة على قيمة الكتلة الاحتمالية لهذا المتغير العشوائي $X$ عند القيمة $x_i$، أي إنَّ $\mathbf{u}(x_i\,;n)\triangleq P(X=x_i)$.

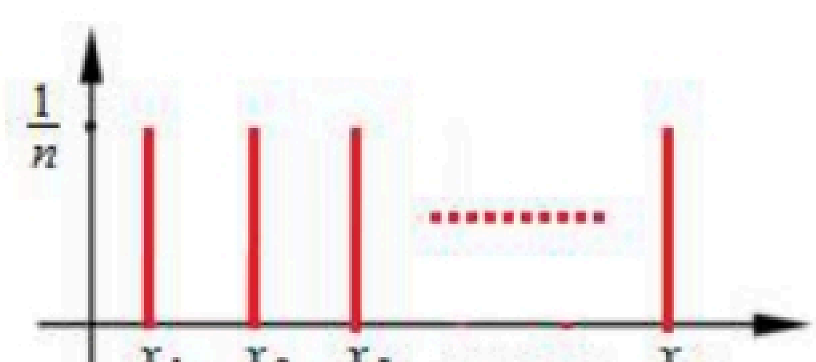
٢- لاحظ هنا أنَّه عندما يكون للمتغيِّر العشوائي $X$ قيمة وحيدة $x_1$ فإنَّ هذا التوزيع يوافق توزيع وحيد النقطة بمَعْلمة $c=x_1$ كحالة خاصّة، وكذلك عندما يكون لـ $X$ قيمتين $x_1$ و $x_2$ فإنَّ هذا التوزيع سيوافق التوزيع الثنائي النقطة بمَعْلمة $p=\frac{1}{2}$ كحالة خاصّة.

٣- يستخدم هذا النوع من المتغيِّرات العشوائية في المسائل التي تقترن فيها أيّة نتيجة من نتائج التجربة العشوائية بإحدى $n$ قيمة حقيقية لها النصيب نفسه في الاختيار.

٤- التمثيل الجدولي لوالبياني لهذا المتغيِّر العشوائي موضَّح كما يلي:

الجدول (٦,٣)

| $i$ | 1 | 2 | ...... | $n$ | sum |
|---|---|---|---|---|---|
| قيم $X$ $x_i$ | $x_1$ | $x_2$ | ...... | $x_n$ | |
| $p_i=P(X=x_i)$ | $\frac{1}{n}$ | $\frac{1}{n}$ | ...... | $\frac{1}{n}$ | 1 |




الشكل (٦,٥)

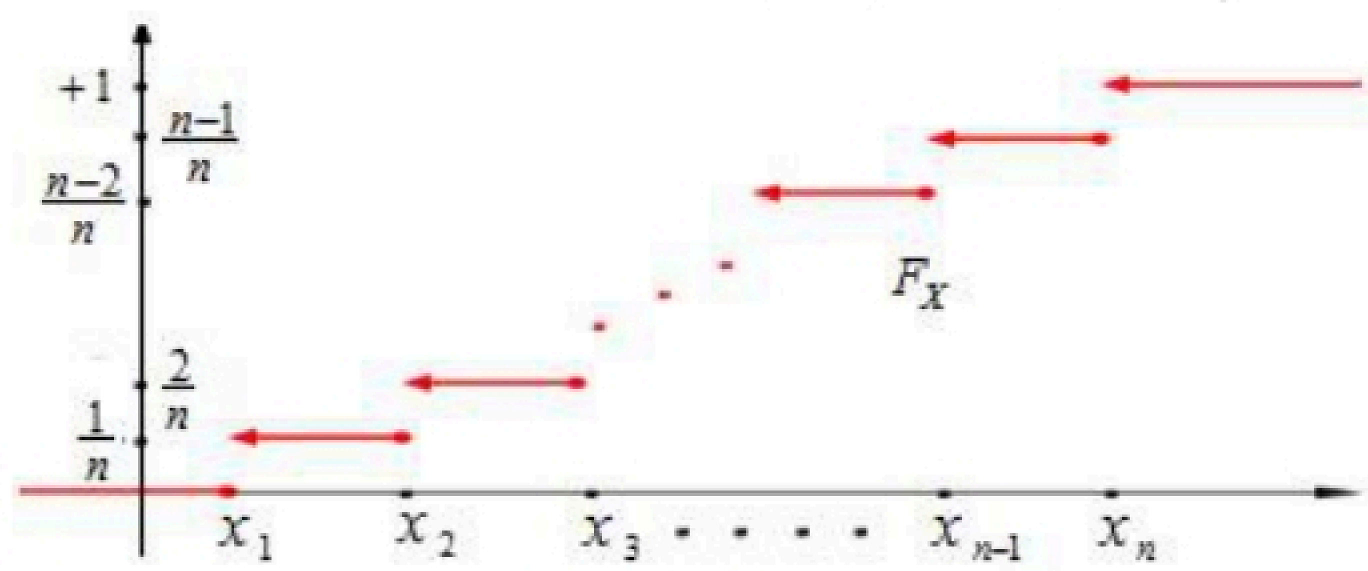
٦- لدى استخدام تعريف دالَّة توزيع متغيِّر عشوائي متقطِّع نجد إنَّ لدالَّة توزيع هذا المتغيِّر العشوائي العرض الاتي:

$$F_X(x)=\frac{i}{n}\,\mathbf{I}_{(x_i,x_{i+1}]}(x)\cdot\mathbf{I}_{\mathbb{N}_{n-1}}(i)+\mathbf{I}_{(x_n,+\infty]}(x)\qquad;\forall x\in\mathbb{R}$$

والتي يمكن كتابتها وفقاً للصيغة الشرطية بالشكل الآتي أيضاً:

$$F_X(x)=\begin{cases} 0 & if\ \ x \le x_1 \\ \dfrac{i}{n} & if\ \ x_i < x \le x_{i+1}\ \ for\ i \in \mathbb{N}_{n-1} \\ 1 & if\ \ x > x_n \end{cases}$$

وأمّا العرض البياني لدالّة توزيع هذا المتغيِّر العشوائي فهي على النحو الآتي:




الشكل (٦,٦)

### (٣,٣,١,٦) أمثلة

١- لنأخذ تجربة إلقاء حجر نرد متوازن لمرة واحدة فقط، وليكن $X$ متغيِّرا عشوائيًّا فوق الفضاء الاحتمالي لهذه التجربة، ومعرفاً من خلال العلاقة الآتية:

$$X(\omega) := \frac{1}{\omega} \qquad ; \forall\ \omega \in \Omega$$

فنجد أنَّ هذا المتغيِّر العشوائي خاضع للتوزيع المنتظم المتقطِّع بمَعْلمة $n = 6$، وأمّا مجموعة قيمه فهي:

$$\mathbf{X} = \left\{ \frac{1}{6}, \frac{1}{5}, \frac{1}{4}, \frac{1}{3}, \frac{1}{2}, 1 \right\}$$

واحتمال أن يأخذ هذا المتغيِّر العشوائي أيّة قيمة من قيمه تساوي $\frac{1}{6}$، وذلك لأنَّه بسبب توازن حجر النرد لدينا:

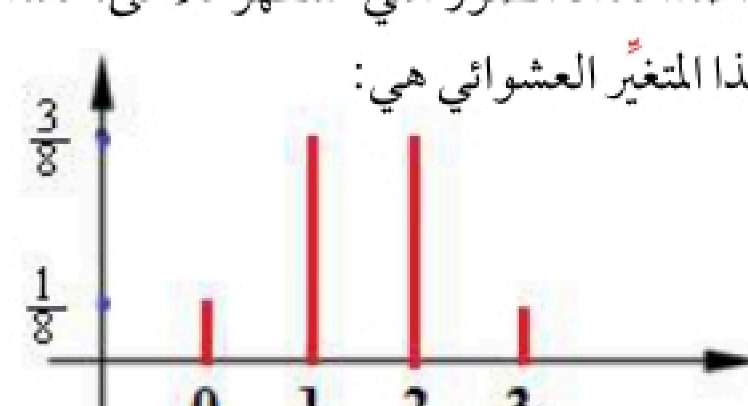
$$P\left(\left\{\omega \in \Omega\, ;\ X(\omega) = \frac{1}{\omega}\right\}\right) = P(\{\omega\}) = \frac{1}{6} \qquad ; \forall\ \omega \in \Omega$$

أي إنَّه من أجل كل $i \in \mathbb{N}_6$ لدينا $P(X = x_i) = \frac{1}{6}$.

٢- لنأخذ تجربة إلقاء قطعة نقود متوازنة لثلاث مرات متتالية، وليكن $X$ متغيِّراً عشوائياً فوق الفضاء الاحتمالي لهذه التجربة راصداً لعدد الصور التي ستظهر للأعلى، فعندئذ نجد أنَّ مجموعة قيمه هي $\mathbf{X} = \{0, 1, 2, 3\}$، وسنفترض أنَّ الاحتمالات الكتلية لهذا المتغيِّر العشوائي هي:

$$P(X = 0) = \frac{1}{8}\ \ \&\ \ P(X = 1) = \frac{3}{8}$$

$$P(X = 2) = \frac{3}{8}\ \ \&\ \ P(X = 3) = \frac{1}{8}$$




الشكل (٦,٦) والتمثيل البياني للمتغيِّر العشوائي $X$

فنلاحظ أنَّ هذا المتغيِّر العشوائي **لا يخضع** للتوزيع المنتظم المتقطِّع، والتمثيل البياني أعلاه لهذا المتغيِّر العشوائي يوضّح لنا ذلك أيضاً.

### (٦, ١, ٤) التوزيع الحدَّاني Binomial Distribution

إنَّ هذا المتغيِّر العشوائي (قدِّم أحد نماذجه كمثال في الفصل السابق) يُعدّ من المتغيِّرات العشوائية المهمَّة في الكثير من الدراسات التطبيقية.

#### (٦, ١, ٤, ١) تعريفه

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، فعندئذ يُقال عن $X$ إنَّه خاضع للتوزيع الحدَّاني بمَعْلمتين $n \in \mathbb{N}$ و $0 < p < 1$ إذا كان يأخذ قيماً صحيحةً غير سالبة، وكان قانون توزيعه مُعطى من خلال العلاقة الآتية:

$$P(X = k) = \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \quad ; k \in \mathbb{N}_n^o \qquad [6,4]$$

نشير هنا إلى أنَّ سبب تسمية هذا التوزيع بهذا الاسم يعود إلى الصيغة التي تُعطي مجموع الكتل الاحتمالية $P(X = k)$ والتي لها صيغة مطابقة للحدّ العام في منشور ثنائي الحد الشهير $(x+y)^n$ والمعروف باسم نشر الكرخي – نيوتن Karkhi - Newton expression (نسبة إلى الفيلسوف والرياضياتي أبو بكر بن محمد بن حسين الكرخي Abū Bakr ibn Muḥammad ibn al Ḥusayn **Al-Karajī** (953-1029)، وكذلك الرياضياتي والفيزيائي الإنكليزي نيوتن Sir Isaac Newton (1642-1726))، ولصيغته هذا النشر العرض الآتي:

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot x^k \cdot y^{n-k}$$

حيث لدينا:

$$1 = (p + (1-p))^n = \sum_{k=0}^{n} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k}$$

#### (٦, ١, ٤, ٢) ملاحظات

١- سوف نستخدم الرمز $\mathbf{B}(n,p)$ للدلالة على هذا التوزيع، وكذلك سنرمز بـ $\mathbf{b}(k;n,p)$ للدلالة على قيمة الكتلة الاحتمالية لهذا المتغير العشوائي $X$ عند القيمة $k$، أي إنَّ:

$$\mathbf{b}(k;n,p) \triangleq P(X = k)$$

٢- إنَّ هذا النوع من المتغيِّرات العشوائية يستخدم في التجارب العشوائية التي هي حصيلة $n$ تكرار مستقل لتجارب برنولية باحتمال نجاح $p$، ويكون في هذه الحالة المتغيِّر العشوائي $X$ راصداً لعدد النجاحات التي نحصل عليها خلال الـ $n$ تجربة. هنا يجب التأكيد على استقلالية التكرارات للتجارب البرنولية، فلو افترضنا على سبيل المثال أنَّ شخصاً ما يقوم بتوزيع بطاقات لعب، فعندئذ نلاحظ أنَّ الاحتمالات للمحاولات المتكرِّرة يتغيَّر إذا لم يتم استبدال البطاقات. فاحتمال سحب بطاقة من نوع القلب في السحب الأول هو 1/4، ولكن في السحب الثاني هو 12/51 أو 13/51، وهذا يتوقف على بطاقة القلب إن كانت قد ظهرت في السحب الأول أم لا، وهنا نكون أمام احتمال شرطي، ومن ثمَّ فلا يمكننا النظر إلى هذه المحاولات المتكرّرة على أنَّها تكرارات برنولية مستقلَّة.

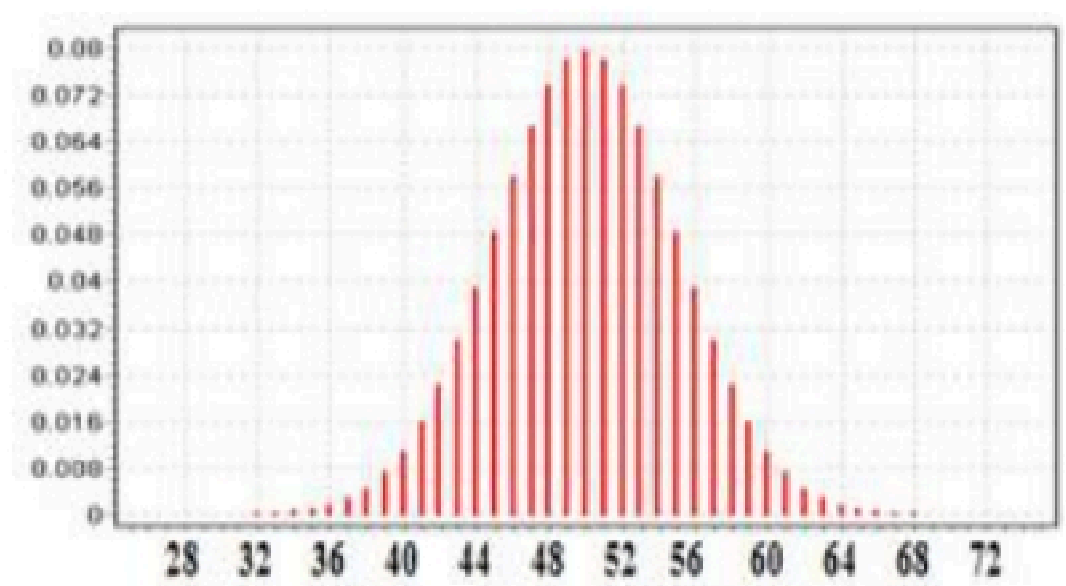
٣- التمثيل الجدولي لهذا المتغيِّر العشوائي (من أجل $n = 5$ و $p = 0.5$) يقدّمه لنا الجدول الآتي:

الجدول (٦, ٤)

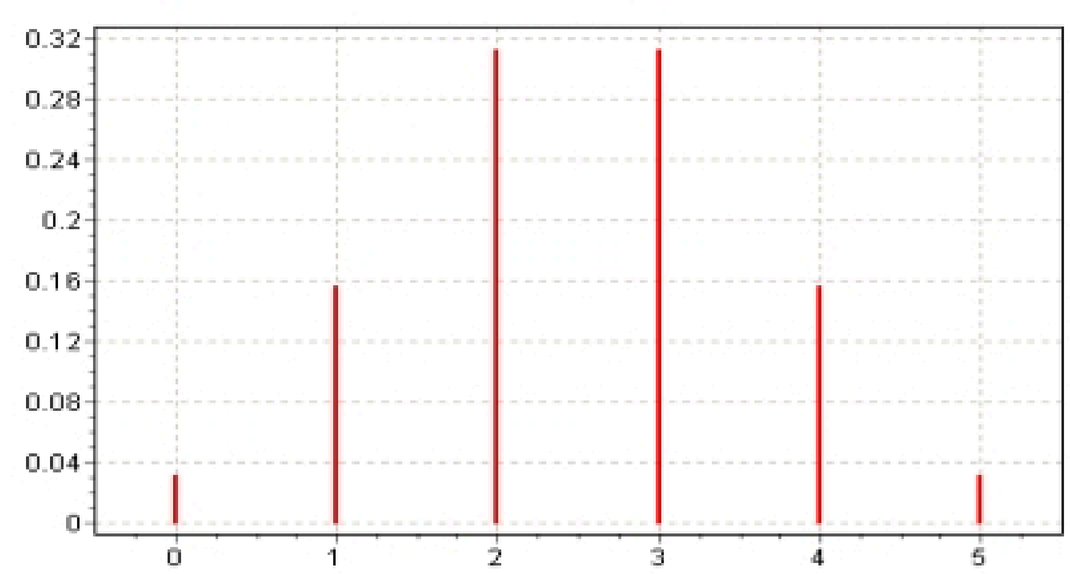
| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | sum |
|---|---|---|---|---|---|---|---|
| قيم $X$ $x_i$ | 0 | 1 | 2 | 3 | 4 | 5 | |
| $p_i = P(X = x_i)$ | 0.031 | 0.156 | 0.313 | 0.313 | 0.156 | 0.031 | 1 |

وأمّا التمثيل البياني لهذا المتغيِّر العشوائي فسوف نقدّمه من خلال عروض عديدة موافقة لقيم مختلفة للمَعْلمتين $n$ و$p$، حيث نحصل

على العروض الآتية:

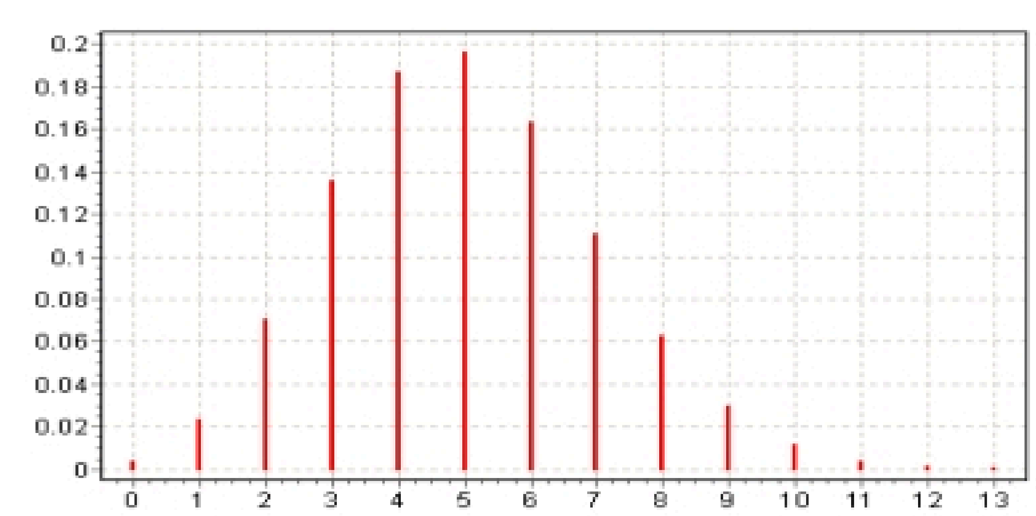



الشكل (٦,٧. أ) من أجل $n = 100$ و $p = 0.5$

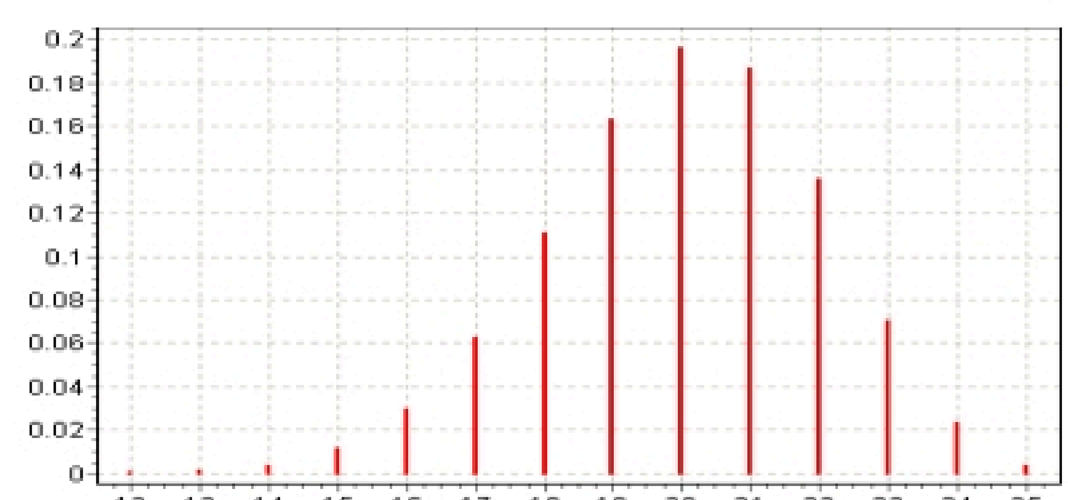



الشكل (٦,٧. ب) من أجل $n = 5$ و $p = 0.5$

فنلاحظ أنَّه من أجل القيم المُعطاة آنفاً سيكون توزيع هذا المتغيّر العشوائي متناظر، وأنَّه في الشكل (٦,٧. أ) لدينا قيم الاحتمالات الموافقة لـ $k < 32$ و $k > 68$ صغيرة لدرجة يصعب إيضاحها بالرسم.




الشكل (٦,٧. ج) من أجل $n = 25$ و $p = 0.2$




الشكل (٦,٧. د) من أجل $n = 25$ و $p = 0.8$

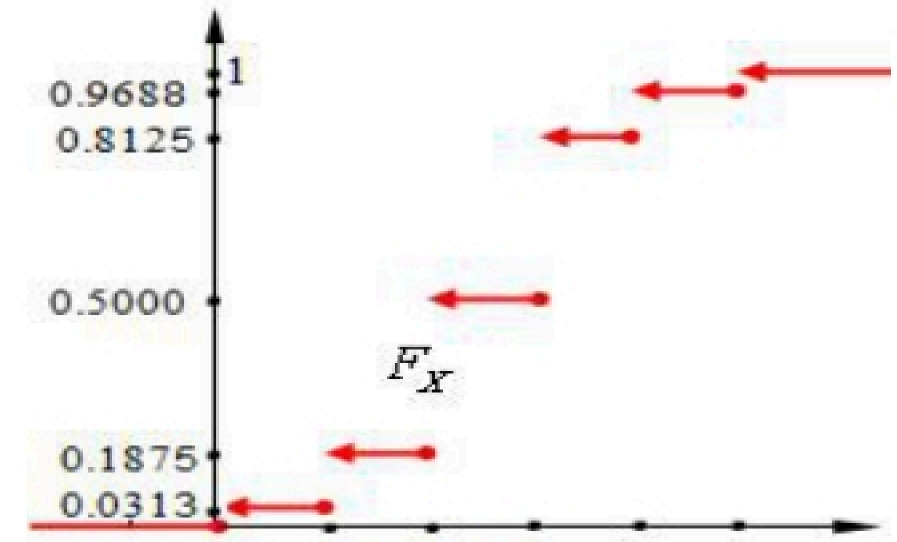
وأمّا هنا، فنلاحظ أنَّه من أجل القيم المُعطاة آنفاً قد أصبح التوزيع ملتوياً، وفي الشكل (٦,٧. ج) الالتواء نحو اليمين وقيم الاحتمالات الموافقة لـ $k > 12$ صغيرة بحيث يصعب إيضاحها بالرسم، وأمّا في الشكل (٦,٧. د) فإنَّنا نجد أنَّ الالتواء نحو اليسار وقيم الاحتمالات الموافقة لـــ $k < 12$ صغيرة بحيث يصعب إيضاحها بالرسم.

٥- إنَّ دالَّة توزيع هذا المتغيِّر العشوائي نجدها باستخدام تعريف دالَّة توزيع متغيِّر عشوائي متقطِّع، حيث لدينا:

$$F_X(x) = \sum_{\substack{k=0 \\ k<x}}^{n} \binom{n}{k} p^k (1-p)^{n-k} \qquad ; \forall x \in \mathbb{R}$$

ومن أجل $n = 5$ و $p = 0.5$ يصبح لها العرض الآتي:

$$F_X(x) = \begin{cases} 0 & for\ x \le 0 \\ 0.0313 & for\ 0 < x \le 1 \\ 0.1875 & for\ 1 < x \le 2 \\ 0.5000 & for\ 2 < x \le 3 \\ 0.8125 & for\ 3 < x \le 4 \\ 0.9688 & for\ 4 < x \le 5 \\ 1 & for\ x > 5 \end{cases}$$




الشكل (٦,٨) دالَّة التوزيع الاحتمالية من أجل $n = 5$ و $p = 0.5$ .

في الحقيقة عندما تكون قيمة المَعْلمة $n$ كبيرةً جداً و$p$ صغيرةً جداً فإنَّ حساب قيمة الدالَّة $F_X$ يصبح شاقاً، لا بل قد يكون من غير الممكن حسابها بالوسائط البسيطة، ولذلك يَلجأ المرء في مثل هذه الحالات إلى استخدام الجداول (كالجدول الخاص بهذا التوزيع الموجود في نهاية هذا الكتاب)، أو استخدام آلات حاسبة متطورة، أو اللجوء إلى استخدام توزيع آخر كتقريب لهذا التوزيع، وسنعالج هذه الفقرة في نهاية هذا الفصل. كما يمكن استخدام صيغ تكاملية مناسبة مثل صيغة تكامل بيتا المنقوص من أجل إعطاء قيم تقريبية مقبولة لقيم هذه الدالَّة، علماً أنَّ الصيغة التكاملية لتكامل بيتا المنقوص تُعطى من خلال العلاقة الآتية:

$$\beta_{n,x} = (n-x)\binom{n}{x}\int_0^q t^{n-x-1}(1-t)^x \, dt \quad ; x \in \mathbb{R}, n \in \mathbb{N}$$

حيث يبرهن أنَّه من أجل كل $x \in \mathbb{R}$ يكون لدينا:

$$F_X(x) = \left(n - \lfloor x \rfloor\right)\binom{n}{\lfloor x \rfloor}\int_0^{1-p} t^{n-\lfloor x \rfloor -1}(1-t)^s \, dt$$

مع العلم أنَّ $\lfloor x \rfloor$ هو أكبر عدد صحيح أصغر أويساوي $x$.

**(٦,١,٤,٣) أمثلة**

١- لنقم بإلقاء قطعة نقود سبع مرات متتالية، فإذا كان احتمال ظهور الشعار في الرمية الواحدة يساوي 0.3، فما هو احتمال الحصول على شعارين على الأقل خلال الرميات السبع؟

**الحـل:** للإجابة على هذا السؤال سنفترض أنَّ $X$ متغيِّر عشوائي راصد لعدد الشعارات التي نحصل عليها خلال الرميات السبع لقطعة النقود، فعندئذ يكون $X$ خاضعاً للتوزيع الحدَّاني بمَعْلمتين $n = 7$ و $p = 0.3$ وذلك لأنَّ التكرارات لعمليّة إلقاء قطعة النقود هي تكرارات لتجارب برنولية مستقلَّة بعضها عن البعض الآخر، واحتمال النجاح في كل منها يساوي $p = 0.3$ ، وأمَّا المطلوب فهو $P(X \geq 2)$ حيث لدينا:

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)] \\ &= 1 - [\mathbf{b}(0;7,0.3) + \mathbf{b}(1;7,0.3)] = 1 - [0.0824 + 0.2471] = 0.6705 \end{aligned}$$

٢- يقوم رجل بالرمي على هدف معيَّن وبحيث لايعلم شيئًا عن إصابته للهدف، فإذا كان احتمال إصابة الرجل للهدف في كل رمية تساوي 0.8، وقام بإطلاق 10 رميات على الهدف، فما هو احتمال أن يصيب الهدف بتسع رميات تماماً؟

**الحـل:** ليكن $X$ متغيِّرًا عشوائيًّا راصدًا لعدد الرميات التي تصيب الهدف، فعندئذ يكون هذا المتغيِّر العشوائي خاضع للتوزيع الحدَّاني بمَعْلمتين $n = 10$ و $p = 0.8$ ؛ وذلك لأنَّ كل عملية رمي من قبل الرجل هي تجربة برنولية بمَعْلمة $p = 0.8$ وتكرارات هذه التجارب مستقل بعضها عن البعض الآخر (لأنَّ الرامي لايعلم شيأ عن إصاباته للهدف)، ومن ثَّم يكون الاحتمال المطلوب هو:

$$P(X = 9) = \binom{10}{9}(0.8)^9(0.2) = 0.2684$$

## (٦,١,٥) التوزيع الحدَّاني السالب Negative Binomial Distribution

إنَّ آلية عمل هذا المتغيِّر العشوائي تشابه آلية عمل المتغيِّر العشوائي السابق، وهو من المتغيِّرات العشوائية المهمَّة أيضاً.

**(٦,١,٥,١) تعريفه**

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ ، فعندئذ يُقال عن $X$ إنَّه خاضع للتوزيع الحدَّاني السالب (أو توزيع باسكال Pascal Distribution نسبة إلى الرياضياتي والفيزيائي الفرنسي باسكال (Blaise Pascal (1623-1662)) بمَعْلمتين $n \in \mathbb{N}$ و $0 < p < 1$ إذا كان يأخذ قيماً صحيحةً غير سالبة، وكان قانون توزيعه مُعطى من خلال العلاقة الآتية:

$$P(X=k)=\binom{k+n-1}{k}\cdot(1-p)^k\cdot p^n \qquad ;k\in\mathbb{N}^o \qquad [6,5\text{-a}]$$

أو وفقاً للصيغة الآتية:

$$P(X=k)=\binom{k+n-1}{n-1}\cdot(1-p)^k\cdot p^n \qquad ;k\in\mathbb{N}^o \qquad [6,5\text{-b}]$$

ونشير هنا إلى أنَّ سبب تسمية هذا التوزيع بهذا الاسم يعود إلى صيغة التوافيق الآتية التي تظهر لنا توافيق لقيم سالبة لـ $n$ :

$$\binom{k+n-1}{k}=(-1)^r\binom{-n}{k}=(-1)^r\frac{(-n)\cdot(-n-1)\cdot\ldots\cdot(-n-k+1)}{(k)\cdot(k-1)\cdot\ldots\cdot(2)\cdot(1)}$$

### (٦,١,٥,٢) ملاحظات

١- سوف نستخدم الرمز $\tilde{\mathbf{B}}(n,p)$ للدلالة على هذا التوزيع، وكذلك سنرمز بـ $\tilde{\mathbf{b}}(k;n,p)$ للدلالة على قيمة الكتلة الاحتمالية لهذا المتغير العشوائي $X$ عند القيمة $k$، أي إنَّ:

$$\tilde{\mathbf{b}}(k;n,p)\triangleq P(X=k)$$

٢- يستخدم هذا النوع من المتغيِّرات العشوائية في التجارب العشوائية التي هي حصيلة تكرارات مستقلّة لتجارب برنولية باحتمال نجاح $p$، وهنا يرصد المتغيِّر العشوائي $X$ عدد التجارب التي نحصل فيها على فشل حتى يتم الحصول على $n$ نجاحاً تماماً.

٣- إذا كان اهتمامنا منصباً على عدد التجارب $k$ التي يجب أن تنفّذ حتى الحصول على $n$ نجاحاً (ومن ثمَّ يكون لدينا $k-n$ فشلاً)، فعندئذ سيصبح لقانون توزيع $X$ العرض الآتي:

$$P(X=k)=\binom{k-1}{n-1}p^n\ (1-p)^{k-n} \qquad [6,5\text{-c}]$$

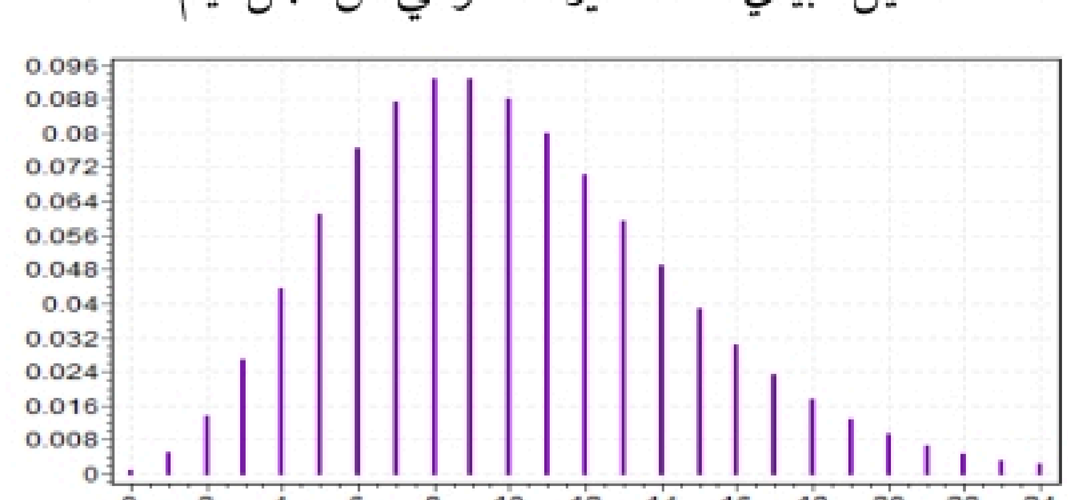
علماً أنَّ قيم $k$ في هذه العلاقة الأخيرة تبدأ من $n+1$، $n+2$ ثمَّ $n+3$ ثمَّ ...

٤- لتوضيح العرض الجدولي لهذا المتغيِّر العشوائي سنأخذ قيماً مُحدَّدة للمَعْلمتين، فمن أجل $n=2$ و $p=0.5$ يكون لدينا التمثيل الجدولي الآتي لهذا المتغيِّر العشوائي.

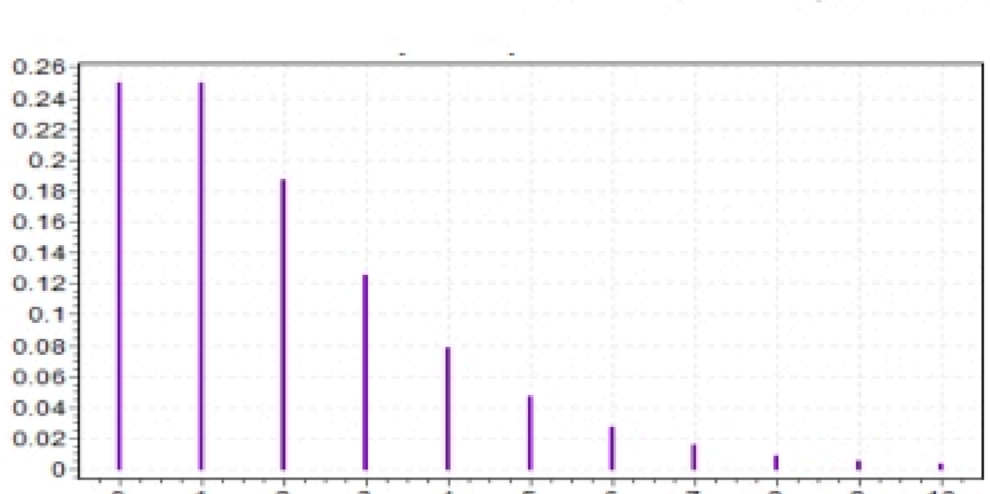
الجدول (٦,٥)

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | sum |
|---|---|---|---|---|---|---|---|
| قيم $X$ $x_i$ | 0 | 1 | 2 | 3 | 4 | 5 | |
| $p_i=P(X=x_i)$ | 0.250 | 0.250 | 0.187 | 0.125 | 0.078 | 0.046 | $\approx 1$ |

٥- التمثيل البياني لهذا المتغيِّر العشوائي من أجل قيم مختلفة لـ $n$ و $p$ يقدّمها الشكلين الآتيين.



الشكل (٩,٦.أ). من أجل $n=10$ و $p=0.5$



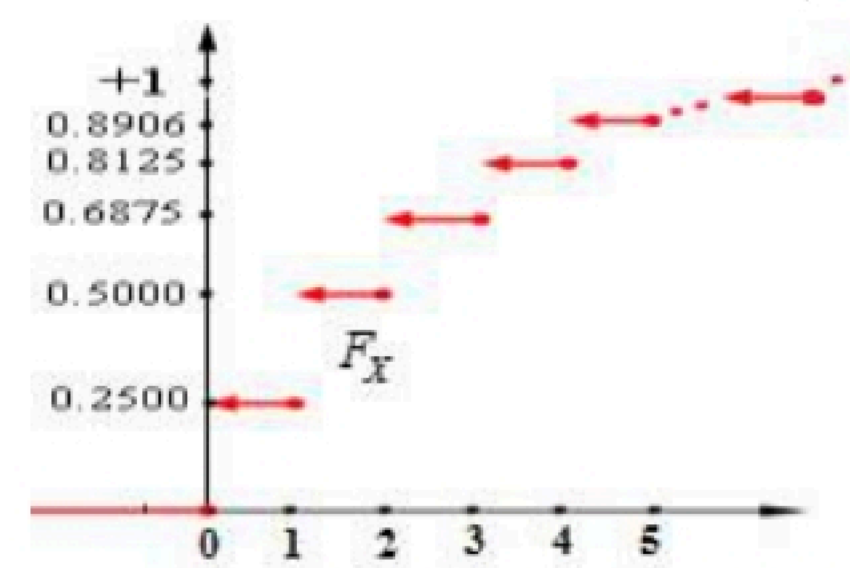
شكل (٩,٦.ب). من أجل $n=2$ و $p=0.5$

٦- باستخدام تعريف متغيِّر عشوائي متقطِّع نجد لدالَّة توزيع هذا المتغيِّر العشوائي العرض الآتي:

$$F_X(x) = p^n \cdot \sum_{\substack{k=0 \\ k<x}} \binom{k+n-1}{k} \cdot (1-p)^k \qquad ; x \in \mathbb{R}$$

والتي يمكن عرضها من أجل $n = 2$ و $p = 0.5$ على النحو الآتي:

$$F_X(x) = \begin{cases} 0 & for \ \ x \le 0 \\ 0.25 & for \ \ 0 < x \le 1 \\ 0.50 & for \ \ 1 < x \le 2 \\ 0.6875 & for \ \ 2 < x \le 3 \\ 0.8125 & for \ \ 3 < x \le 4 \\ \vdots & \vdots \ \ \vdots \ \ \ \vdots \ \ \vdots \\ 1 & for \ \ x \to \infty \end{cases}$$




الشكل (٦,١٠). دالّة التوزيع الاحتمالية من أجل $n = 2$ و $p = 0.5$ .

### (٣,٥,١,٦) أمثلة

١- لنفترض أنَّ باحثاً يقوم بتنفيذ تجربة للتحقُّق من نتيجة معيَّنة، ويكرِّر تنفيذ هذه التجربة تحت الشروط نفسها وبشكل مستقل كل منها عن الأخرى، فإذا افترضنا أنَّ احتمال تحقُّق النتيجة المرجوة يساوي 0.05، وأنَّه سيتوقف عن تنفيذ التجربة عندما تتحقَّق النتائج المرجوة في ثلاث منها، فما هو احتمال أن يقوم الباحث بتنفيذ خمس تجارب تماماً؟

الحـل: في هذا المثال نلاحظ أنَّ التجارب الخمس هي تجارب برنولية مستقلّ بعضها عن بعض، واحتمال النجاح لكل تجربة هو $p = 0.05$ ، فلو كان $X$ متغيِّر عشوائي راصد لعدد التجارب التي سيفشل الباحث في تنفيذها حتى يتمّ الحصول على ثلاث نتائج إيجابية بخصوص المسألة التي هي قيد الدرس، لوجدنا أنَّ هذا المتغيِّر العشوائي خاضع للتوزيع الحدَّاني السالب بمَعْلمتين $n = 3$ و $p = 0.05$ ، ومن ثَمَّ احتمال الحصول على فشلين فقط قبل حصولنا على النجاح الثالث يساوي:

$$P(X = 2) = \bar{\mathbf{b}}(2;3,0.05) = \binom{2+3-1}{2}(0.05)^3(0.95)^2 = 0.0007$$

٢- تقوم آلة بانتاج نوع من السجاد، فإذا افترضنا أنَّ احتمال إنتاج سجادة خالية من العيوب في كل مرّة يساوي 0.85 وأنَّ الآلة ستتوقف عن الإنتاج إذا صار عدد السجاد المعيب في إنتاجها يساوي ثلاث سجادات، فما هو احتمال توقف الآلة عن الإنتاج بعد تصنيع السجادة التاسعة؟

الحـل: في هذا المثال نلاحظ أنَّ عملية إنتاج السجاد هي تجارب برنولية مستقلّ بعضها عن البعض الآخر باحتمال نجاح $p = 0.15$ (لأنَّ النجاح هنا هو الحصول على سجادة معيبة)، فلو كان $X$ متغيِّر عشوائي راصد لعدد السجادات السليمة حتى الحصول على ثلاث سجادات معيبة، فإنَّنا نجد أنَّ هذا المتغيِّر العشوائي خاضع للتوزيع الحدَّاني السالب بمَعْلمتين $n = 3$ و $p = 0.15$ ، ومن ثَمَّ الاحتمال المطلوب هو:

$$P(X = 6) = \bar{\mathbf{b}}(6;3,0.15) = \binom{6+3-1}{6}(0.15)^3(0.85)^6 = 0.0356$$

٣- يقوم شخص ما بقذف قطعة نقود متوازنة، فما هو احتمال أن يقوم بتكرار عمليّة القذف لتسع مرّات حتى يحصل في المرَّة الأخيرة على الصورة الخامسة؟

الحـل: ليكن $X$ متغيِّرا عشوائيًّا راصدًا لعدد الصور التي سيحصل عليها لدى تنفيذ عملية القذف، فعندئذ بسبب أنَّ عملية قذف قطعة النقود هي تجارب برنولية مستقلّة بعضها عن البعض الآخر باحتمال نجاح $p = 0.5$ ، فإنَّنا نجد لهذا المتغيِّر العشوائي $X$ توزيعًا حدَّانيًّا سالبًا بمَعْلمتين $n = 5$ و $p = 0.5$ ، ومن ثَمَّ الاحتمال المطلوب هو:

$$P(X = 9) = \tilde{\mathbf{b}}(9;5,0.5) = \binom{9-1}{5-1} 0.5^5 \ (1-0.5)^{9-5} = 0.1367$$

٤ – لدينا مستودع لإطارات (كفرات) السيارات جميعها مُغلَّفة، وعادة يكون من بين هذه الإطارات كميَّة من الإطارات المعيبة (غير مطابقة للمواصفات) ولكنَّها غير مرئية، فإذا افترضنا أنَّ نسبة الإطارات المعيبة في هذا المستودع تساوي 0.05، وأردنا أن نأخذ أربعة إطارات لسيّارة، فما هو احتمال عثورنا على إطارين معيبين قبل الحصول على الإطارات الأربع السليمة؟

الحـل: في هذا المثال نلاحظ أنَّ البحث عن إطارات للسيّارة هي تجارب برنولية مستقلّة بعضها عن البعض الآخر باحتمال نجاح $p = 0.95$ (لأنَّ النجاح هنا هو الحصول على إطار سليم)، فلو أخذنا $X$ متغيِّرا عشوائيًّا راصدًا لعدد الإطارات السليمة حتى الحصول على إطارين معيبين، فإنَّنا نجد أنَّ هذا المتغيِّر العشوائي خاضع للتوزيع الحدَّاني السالب بمَعْلمتين $n = 4$ و $p = 0.95$ ، ومن ثَمَّ الاحتمال المطلوب هو:

$$P(X = 2) = \tilde{\mathbf{b}}(2;4,0.95) = \binom{2+4-1}{2} (0.95)^4 (0.05)^2 = 0.0204$$

### (٦, ١, ٦) التوزيع الهندسي Geometric Distribution

إنَّ آليّة عمل هذا المتغيِّر العشوائي تنتج عن آليّة عمل المتغيِّر العشوائي السّابق في حالة خاصّة، وهو من المتغيِّرات العشوائية المهمَّة أيضاً، ويقدِّمه لنا التعريف الآتي.

#### (٦, ١, ٦, ١) تعريفه

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ ، فعندئذٍ يُقال عن $X$ إنَّه خاضع للتوزيع الهندسي بمَعْلمة $0 < p < 1$ إذا كان يأخذ قيماً صحيحةً موجبةً، وكان قانون توزيعه مُعطى من خلال العلاقة الآتية:

$$P(X = k) = p \cdot (1-p)^{k-1} \qquad ; k \in \mathbb{N} \qquad [6,6]$$

ونشير هنا إلى أنَّ سبب تسمية هذا التوزيع بهذا الاسم يعود إلى نموذج الدالَّة المولِّدة الاحتمالية لهذا التوزيع حيث لها صيغة دالَّة هندسية (راجع فصل العزوم للاطلاع على مفهوم الدالَّة المولِّدة الاحتمالية).

#### (٦, ١, ٦, ٢) ملاحظات

١ – سوف نستخدم الرمز $\mathbf{G}(p)$ للدلالة على هذا التوزيع، وكذلك سنرمز بـ $\mathbf{g}(k;p)$ للدلالة على قيمة الكتلة الاحتمالية لهذا المتغير العشوائي $X$ ، أي إنَّ:

$$\mathbf{g}(k;p) \triangleq P(X = k)$$

٢ – إنَّ العلاقة الآتية تعدّ من العلاقات المميِّزة لهذا التوزيع:

$$P(X \geq \ell \mid X \geq k) = P(X \geq \ell - k) \ ; \ \ell, k \in \mathbb{N}, \ell > k$$

ومعنى هذه العلاقة أنَّ التوزيع الهندسي لا يمتلك ذاكرة Memoryless، بمعنى أنَّ التوزيع الهندسي ينسى ما قد حصل له في الماضي، ولهذه العلاقة تطبيقات مهمَّة في مجال نظرية الخدمة وقطاعات عشوائية أخرى، وأمّا إثباتها فيتمّ على النحو الآتي:

من كون الحادث $\{\omega \in \Omega ; X(\omega) \geq k\}$ محتوى في الحادث $\{\omega \in \Omega ; X(\omega) \geq \ell\}$، وباستخدام مفهوم الاحتمال الشرطي،

وملاحظة أنَّ الاحتمال $P(X \geq s)$ يعني احتمال عدم حصولنا على أي نجاح خلال التكرارات المستقلّة التي عددها $s$ والذي قيمته تساوي $(1-p)^s$، فإنَّه يمكننا أن نكتب ما يلي:

$$P(X \geq \ell \mid X \geq k) = \frac{P(X \geq \ell, X \geq k)}{P(X \geq k)} = \frac{P(X \geq \ell)}{P(X \geq k)} = \frac{(1-p)^{\ell-1}}{(1-p)^{k-1}} = (1-p)^{\ell-k} = P(X \geq \ell - k)$$

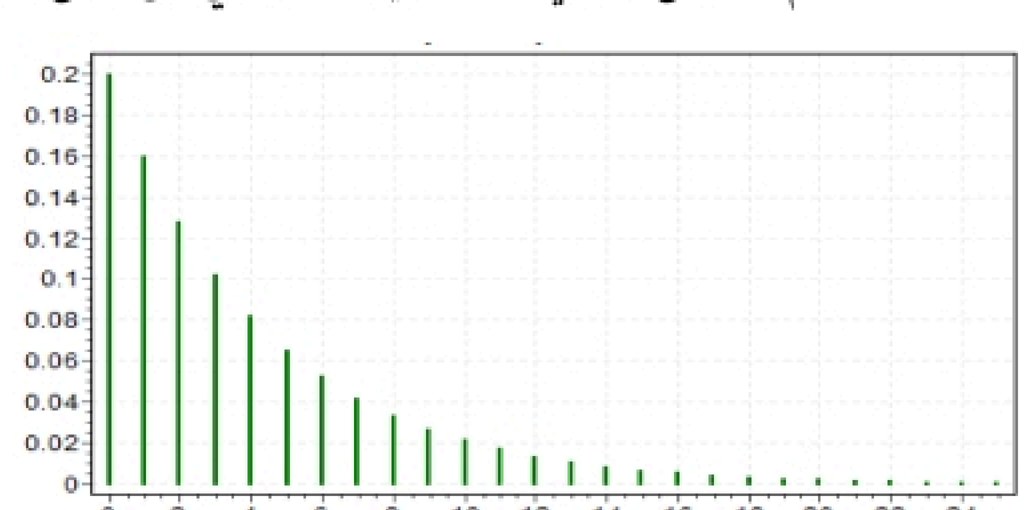
٣- يستخدم هذا النوع من المتغيِّرات العشوائية في التجارب العشوائية التي هي حصيلة تكرارات مستقلّة لتجارب برنولية باحتمال نجاح $p$، علماً أنَّ المتغيِّر العشوائي $X$ هنا يرصد عدد التجارب التي يجب تنفيذها حتى نحصل على أول نجاح. لاحظ هنا أنَّ التوزيع الهندسي يُمثِّل حالة خاصّة من التوزيع الحدَّاني السالب عندما تكون قيمة مَعْلمته $n = 1$ .

٤- الجدولي الاتي يقدِّم لنا هذا المتغيِّر العشوائي من أجل قيمة محدَّدة للمَعْلمة $p$، (سنأخذ $p = 0.5$ ).

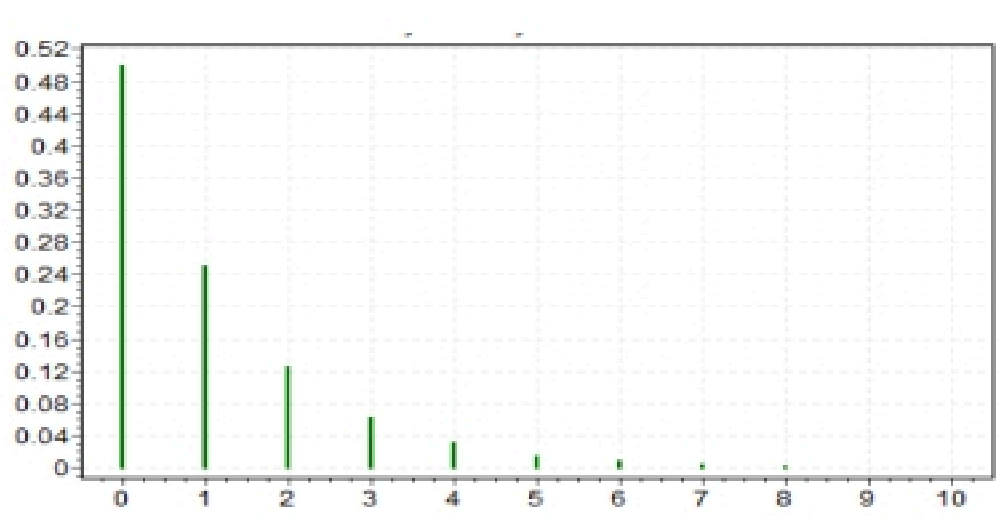
الجدول (٦,٦)

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | sum |
|---|---|---|---|---|---|---|---|
| قيم $X$ $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 | |
| $p_i = P(X = x_i)$ | 0.500 | 0.250 | 0.125 | 0.063 | 0.031 | 0.016 | $\approx 1$ |

٥- سنقدِّم التمثيل البياني لهذا المتغيِّر العشوائي من أجل قيمتين مختلفتين لـ $p$، ويقدّمها الشكلان الآتيان:



الشكل (١١ ,٦ . أ) من أجل $p = 0.2$



شكل (١١ ,٦ . ب) من أجل $p = 0.5$

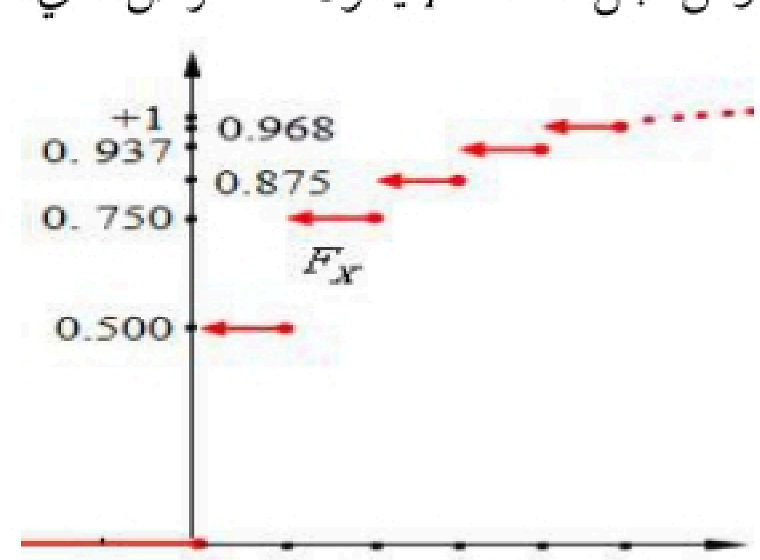
فنلاحظ هنا أنَّ التوزيع ملتوي نحو اليمين، وأنَّه في الشكل (١١ ,٦ . ب) نجد قيم الاحتمالات الموافقة لـ $k > 10$ صغيرة بحيث يصعب إيضاحها بالرسم.

٦- إنَّ دالَّة توزيع هذا المتغيِّر العشوائي نجدها باستخدام تعريف دالَّة توزيع متغيِّر عشوائي متقطِّع، حيث لدينا:

$$F_X(x) = \sum_{\substack{k=1 \\ k<x}} p \cdot (1-p)^{k-1} \qquad ; x \in \mathbb{R}$$

ومن أجل $p = 0.5$ يكون لها العرض الآتي:

$$F_X(x) = \begin{cases} 0 & for\ x \leq 1 \\ 0.500 & for\ 1 < x \leq 2 \\ 0.750 & for\ 2 < x \leq 3 \\ 0.875 & for\ 3 < x \leq 4 \\ \vdots & \vdots \quad \vdots \quad \vdots \quad \vdots \\ 1 & for\ x \to \infty \end{cases}$$




الشكل (١٢ ,٦) دالَّة التوزيع الاحتمالية من أجل $p = 0.5$ .

**(٦,١,٦,٣) أمثلة**

١- لدينا صندوق يحوي $n$ كرة متماثلة تماماً وسجّل على إحداها كلمة فائز، فلو قام شخص ما بعمليّات سحب عشوائية لكرة من هذا الصندوق، ثم أعادها إلى الصندوق بحيث يتوقف عن تكرار عمليّة السحب لدى حصوله على الكرة التي كتب عليها فائز، فما هو احتمال أن يتوقف عن السحب بعد تنفيذ $\frac{n}{2}$ من عمليّات السحب؟

الحـل: في هذا المثال نلاحظ أنَّ عمليّات السحب هي تجارب برنولية مستقلّة بعضها عن البعض الآخر باحتمال نجاح $p=\frac{1}{n}$ ، فلو كان $X$ متغيّراً عشوائياً راصداً لعدد التجارب التي يجب تنفيذها حتى يحصل على الكرة التي كُتِبَ عليها كلمة فائز، لوجدنا أنَّ هذا المتغيِّر العشوائي خاضع للتوزيع الهندسي بمَعْلمة $p=\frac{1}{n}$ ، ومن ثمَّ يكون لدينا:

$$P\left(X=\frac{n}{2}\right)=\frac{1}{n}\left(1-\frac{1}{n}\right)^{\frac{n}{2}-1}$$

فلو أخذنا على سبيل المثال $n=50$ لوجدنا أنَّ احتمال الحصول على الكرة التي كتب عليها كلمة فائز بعد 25 عمليّة سحب يساوي:

$$P(X=25)=\mathbf{g}(25;0.02)=\frac{1}{50}(1-0.02)^{24}=0.0123$$

بينما لو أخذنا $n=100$ لوجدنا أنَّ احتمال الحصول على الكرة المذكورة آنفاً بعد 50 عمليّة سحب يساوي:

$$P(X=50)=\mathbf{g}(50;0.01)=\frac{1}{100}(1-0.01)^{49}=0.00061$$

حيث نلاحظ تضاؤل قيمة هذا الاحتمال كلّما كَبُر عدد الكرات الموجودة في الصندوق.

٢- يقوم شخص ما بقذف قطعة نقود متوازنة حتى يحصل على صورة فيتوقف عن قذف القطعة، فعندئذ:

أ- ما هو احتمال أن يحصل على الصورة لأول مرَّة في عمليّة القذف السّابعة؟

ب- إذا كان الشخص لم يحصل على صورة حتى القذفة الخامسة، فما هو احتمال أن يتابع في قذف القطعة لسبع مرَّات على الأقل؟

الحـل: من أجل الطلب:

أ- ليكن $X$ متغيّرا عشوائيًّا راصدًا لعدد عمليّات قذف قطعة النقود حتى الحصول على الصور للمرّة الأولى، فعندئذ بسبب أنَّ عملية قذف قطعة النقود هي تجارب برنولية مستقلّ بعضها عن بعض باحتمال نجاح $p=0.5$ ، فإنَّنا نجد أنَّ لهذا المتغيِّر العشوائي $X$ توزيعًا هندسيًّا بمَعْلمة $p=0.5$ ، ومن ثمَّ الاحتمال المطلوب هو:

$$P(X=7)=\mathbf{g}(7;0.5)=0.5\,(1-0.5)^6=0.0078$$

ب- إنَّ الاحتمال المطلوب هو $P(X>7\,|\,X>5)$ ، ومن ثمَّ يكون لدينا:

$$P(X>7\,|\,X>5)=P(X>7-5)=1-P(X\le 2)=1-\left[P(X=1)+P(X=2)\right]$$

$$=1-\left[\mathbf{g}(1;0.5)+\mathbf{g}(2;0.5)\right]=1-(0.5+0.25)=0.25$$

ومن هذه النتيجة نلاحظ وضوح معنى أنَّ هذا التوزيع لا يمتلك ذاكرة، وذلك لأنَّ قيمة الاحتمال الذي حصلنا عليه يساوي أيضاً:

$$P(X=2)=0.5\,(1-0.5)=0.25$$

وكأنَّ الشخص قد بدأ بقذف قطعة النقود لأول مرَّة بعد المرَّة الخامسة، وما كان من قذفات سابقة لا تأثير لها على قيمة الاحتمال، وكأنَّ القذفات الخمس الأولى لم تنفَّذ أصلاً (جاءت تسمية هذا التوزيع: بيت من ذاكرة التوزيع).

٣- لدينا جهاز يقوم بإجراء اختبارات لعيّنات صناعية وبشكلٍ مستقل كل منها عن الآخر، فإذا كان احتمال إعطاء نتائج خاطئة

من قبل الجهاز 0.005، وأنَّ الجهاز سيوقف عن العمل عند الحصول على أول نتيجة خاطئة، فعندئذ:

أ - ما هو احتمال أن يتوقف الجهاز عن العمل بعد الانتهاء من اختبار العيّنة ذات الرقم 70؟

ب- إذا كان الجهاز لم يرتكب أي خطأ حتى الاختبار الخامس، فما هو احتمال أن يتابع في عمله حتى الاختبار التاسع على الأقل؟

الحـل: من أجل الطلب:

أ- في هذا المثال نلاحظ أنَّ عملية اختبار العيّنات التي يقوم بها الجهاز هي تجارب برنولية مستقلّ بعضها عن بعض باحتمال نجاح $p = 0.005$ ، حيث النجاح هنا هو حصولنا على نتيجة خاطئة من قبل الجهاز، فلو كان $X$ متغيِّراً عشوائياً راصداً لعدد التجارب التي سيقوم الجهاز بتنفيذها حتى يقع الخطأ الأول من قبل الجهاز، فإنّنا نلاحظ أنَّ هذا المتغيِّر العشوائي خاضع للتوزيع الهندسي بمَعْلمة $p = 0.005$ ، ومن ثمَّ الاحتمال المطلوب يساوي:

$$P(X = 70) = \mathbf{g}(70;0.005) = (0.005)(0.995)^{69} = 0.00354$$

ب- في هذه الحالة لدينا الاحتمال المطلوب هو $P(X > 9 | X > 5)$ ، ومن ثمَّ يكون لدينا:

$$\begin{aligned}
P(X > 9 | X > 5) &= P(X > 9-5) = 1-P(X \le 4) \\
&= 1-\left[P(X=1)+P(X=2)+P(X=3)+P(X=4)\right] \\
&= 1-\left[\mathbf{g}(1;0.005)+\mathbf{g}(2;0.005)+\mathbf{g}(3;0.005)+\mathbf{g}(4;0.005)\right] \\
&= 1-\left[0.005+0.004975+0.00495+0.00493\right] = 0.80145
\end{aligned}$$

٤ - لنفترض أنَّ احتمال نجاح طالب في اختبار تحريري لمقرّر معيَّن يساوي 0.75، وأنَّ الاختبارات يمكن تكرارها بشكل مستقل كل منها عن الآخر (بمعنى أنّه لا يمكن للطالب أن يخمّن طبيعة الأسئلة التي يمكن أن تقدَّم في الاختبار التالي للمقرَّر)، ولنقم بحساب:

أ - احتمال أن ينجح من الاختبار الأول.

ب- احتمال أن يتقدم الطالب للاختبار ثلاث مرّات حتى يُحقِّق النجاح في هذا المقرّر.

ج- احتمال أن يتقدم الطالب للاختبار لثلاث مرّات على الأقل حتى يُحقِّق النجاح في هذا المقرّر.

الحـل: ليكن $X$ متغيِّراً عشوائياً راصداً لعدد الاختبارات التي يتقدَّم بها الطالب حتى يُحقِّق النجاح في المقرَّر، فعندئذ بسبب أنَّ عمليّة تكرار الاختبارات هي تجارب برنوليّة مستقلّ بعضها عن بعض باحتمال نجاح 0.75 ، فإنّنا نجد أنَّ لهذا المتغيِّر العشوائي $X$ توزيعًا هندسيًّا بمَعْلمة $p = 0.75$ ، ومن ثمَّ يكون لدينا من أجل الطلب:

أ- الاحتمال المطلوب هو:

$$P(X = 1) = \mathbf{g}(1;0.75) = 0.75\,(1-0.75)^0 = 0.75$$

ب- الاحتمال المطلوب هو:

$$P(X = 3) = \mathbf{g}(3;0.75) = 0.75\,(1-0.75)^2 = 0.047$$

ج- الاحتمال المطلوب هو:

$$\begin{aligned}
P(X \ge 3) &= 1-P(X \le 2) = 1-\left[P(X=1)+P(X=2)\right] \\
&= 1-\left[\mathbf{g}(1;0.75)+\mathbf{g}(2;0.75)\right] \\
&= 1-\left[0.75\,(1-0.75)^0+0.75\,(1-0.75)\right] = 0.0625
\end{aligned}$$

### (٦,١,٧) التوزيع فوق الهندسي (أو الفوهندسي) Hypergeometric Distribution:

إنَّ آليّة عمل هذا المتغيِّر العشوائي تختلف جذرياً عن آليّة عمل المتغيِّرات العشوائي السابقة، وهو من المتغيِّرات العشوائية المهمَّة في مجال ضبط الجودة الإحصائي.

#### (٦,١,٧,١) تعريفه

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فعندئذ يُقال عن $X$ إنَّه خاضع للتوزيع **فوق الهندسي** بمَعَالم $n$، $M$ و $N \in \mathbb{N}$ مع $N \geq n$ و $N > M$، إذا كان $X$ يأخذ قيماً صحيحةً غير سالبة، وقانون توزيعه مُعطى من خلال العلاقة الآتية:

$$P(X=k)=\binom{M}{k}\cdot\binom{N-M}{n-k}\cdot\binom{N}{n}^{-1} \qquad [6,7]$$

علماً أنَّ $\max\{0,\, n-M+N\} \leq k \leq \min\{n,M\}$.

نشير هنا إلى أنَّ سبب تسميّة هذا التوزيع بهذا الاسم يعود إلى نموذج الدالَّة المولِّدة الاحتمالية لهذا التوزيع حيث لها صيغة دالَّة فوق هندسيّة.

#### (٦,١,٧,٢) ملاحظة

١- سوف نستخدم الرمز $\mathbf{H}(N,M,n)$ للدلالة على توزيع هذا المتغير العشوائي، وكذلك سنرمز بـ $\mathbf{h}(k;N,M,n)$ للدلالة على قيمة الكتل الاحتمالية لـ $X$، أي إنَّ $\mathbf{h}(k;N,M,n) \triangleq P(X=k)$.

٢- يستخدم هذا النوع من المتغيِّرات العشوائية في اختبارات ضبط الجودة الإحصائي وللقيم $n$، $M$، $N$ و $k$ الأدوار الآتية:

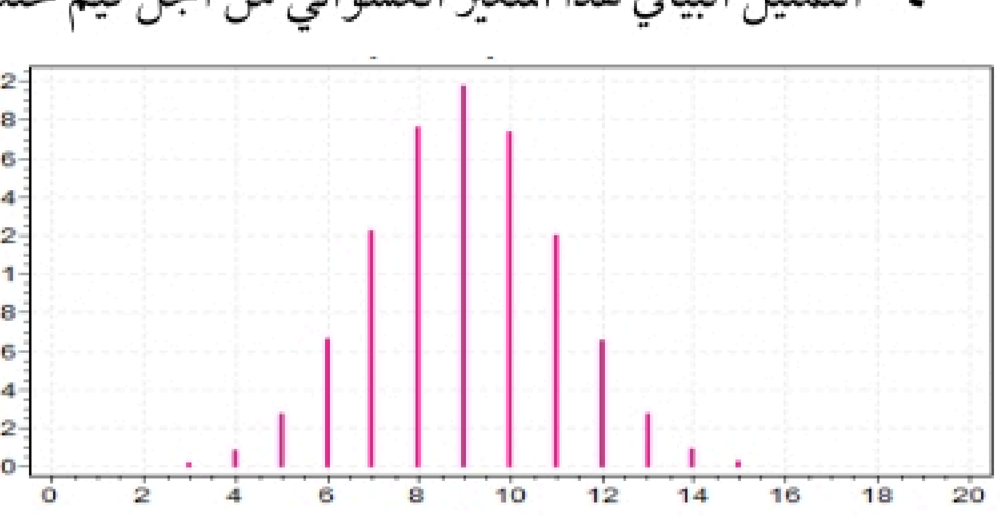
- القيمة $N$ تمثّل حجم المجتمع الإحصائي.
- القيمة $M$ تمثِّل عدد النجاحات في المجتمع الإحصائي.
- القيمة $n$ تمثّل حجم العيّنة.
- القيمة $k$ تمثِّل عدد النجاحات في العيّنة.

٣- لتوضيح العرض الجدولي لهذا المتغيِّر العشوائي سنأخذ قيماً محدَّدة للمَعَالم، فمن أجل $N=10$، $M=5$ و $n=5$ يكون لجدول توزيع هذا المتغيِّر العشوائي العرض الآتي:

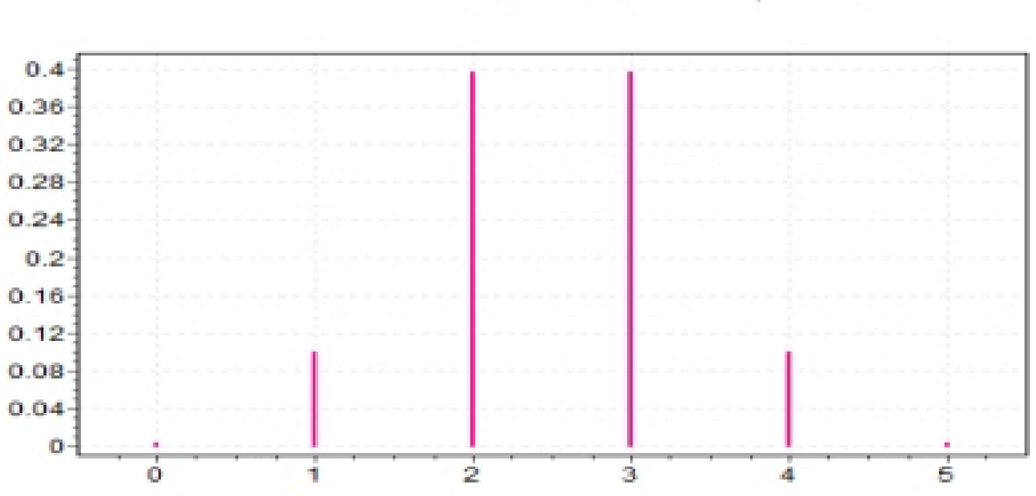
الجدول (٦,٧).

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | sum |
|---|---|---|---|---|---|---|---|
| قيم $X$ $x_i$ | 0 | 1 | 2 | 3 | 4 | 5 | |
| $p_i=P(X=x_i)$ | 0.004 | 0.099 | 0.397 | 0.397 | 0.099 | 0.004 | 1 |

٤- التمثيل البياني لهذا المتغيِّر العشوائي من أجل قيم مختلفة لـــ $N$، $M$ و $n$ يقدّمها الشكلين الآتيين:



الشكل (٦,١٣.أ) من أجل $N=100$ و $M=20$ و $n=45$

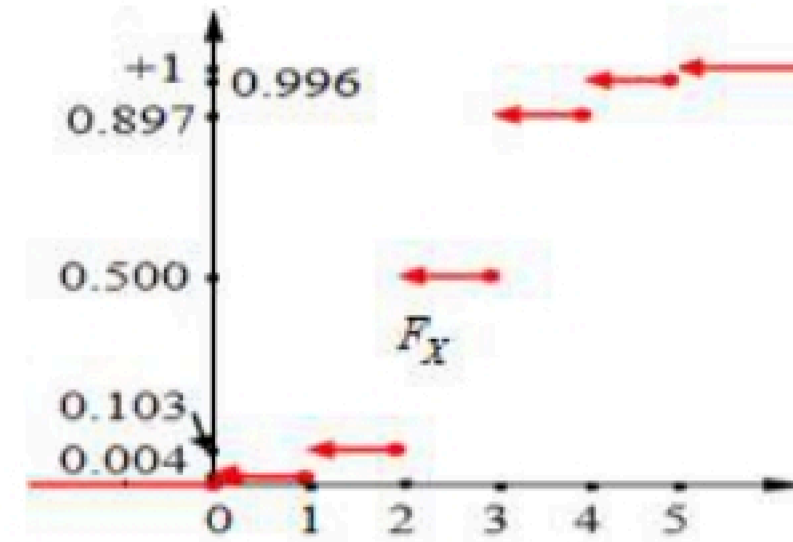


الشكل (٦,١٣.ب) من أجل $N=10$ و $M=5$ و $n=5$

٥- إنَّ دالَّة توزيع هذا المتغيِّر العشوائي نجدها باستخدام تعريف دالَّة توزيع متغيِّر عشوائي متقطِّع، حيث لدينا:

$$F_X(x) = \sum_{\substack{k=0 \\ k<x}} \binom{M}{k} \cdot \binom{N-M}{n-k} \cdot \binom{N}{n}^{-1} \quad ; \forall x \in \mathbb{R}$$

ومن أجل $N = 10$ و $M = 5$ و $n = 5$ يكون لها العرض الآتي:




$$F_X(x) = \begin{cases} 0 & for \;\; x \le 0 \\ 0.004 & for \;\; 0 < x \le 1 \\ 0.103 & for \;\; 1 < x \le 2 \\ 0.500 & for \;\; 2 < x \le 3 \\ 0.897 & for \;\; 3 < x \le 4 \\ 0.996 & for \;\; 4 < x \le 5 \\ 1 & for \;\; x > 5 \end{cases}$$

الشكل (٦,١٤) دالَّة التوزيع الاحتمالية من أجل $N = 10$ و $M = 5$ و $n = 5$.

حيث نلاحظ صعوبة تعيين دالَّة التوزيع $F_X$ من أجل القيم الكبيرة للوسطاء $N$، $M$ و $n$. في الحقيقة تُعدّ العلاقة السابقة لدالَّة التوزيع الاحتمالية لهذا المتغيِّر العشوائي هي أبسط صيغة لحساب الدالَّة $F_X$، ولكن بُذلت محاولات جادة في إيجاد صيغ تقريبية لحساب $F_X$ اعتماداً على توزيعات متقطِّعة أو مستمرة، حيث سنستعرض بعضها في نهاية هذا الفصل (انظر تقارب التوزيعات).

### (٦,١,٧,٣) أمثلة

١- لدينا في جهاز إلكتروني يحوي عشرة دارات متكاملة منها دارتين مُعطَّلتين، قام أحد عمال الصيانة وبشكل عشوائي بإزالة ثلاث دارات من الجهاز، فإذا افترضنا أنَّ لجميع الدارات النصيب نفسه في الاختيار، فما هو احتمال أن تكون الداراتُ المعطَّلة من بين الدارات التي تمّ إزالتها من الجهاز؟

الحـل: بفرض أنَّ $X$ متغيِّرًا عشوائيًّا راصدًا لعدد الدارات المُعطَّلة التي نحصل عليها نتيجة لعمليّة إزالة الدارات، فإنَّنا نجد أنَّ هذا المتغيِّر العشوائي خاضع للتوزيع فوق الهندسي بمَعَالم $N = 10$ و $M = 2$ و $n = 3$، ومن ثمَّ يكون الاحتمال المطلوب هو:

$$P(X = 2) = \mathbf{h}(2;10,2,3) = \binom{2}{2} \cdot \binom{8}{1} \cdot \binom{10}{3}^{-1} = 0.0667$$

٢- في مزرعة إبل يوجد 25 رأساً منها خمسة مغاتير. قام صاحب المزرعة بأخذ عشرة منها عشوائياً وبيعها، فما هو احتمال أن يكون هناك ثلاثة على الأكثر من المغاتير قد تمّ بيعها؟

الحـل: ليكن $X$ متغيِّراً عشوائياً راصداً لعدد المغاتير التي مع الإبل المُباعة، فعندئذ نجد أنَّ هذا المتغيِّر العشـوائي خاضع للتوزيع فوق الهندسي بمَعَالم $N = 25$ و $M = 5$ و $n = 10$، وأمّا الاحتمال المطلوب فهو $P(X \le 3)$ حيث لدينا:

$$\begin{aligned} P(X \le 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= \mathbf{h}(0;10,2,3) + \mathbf{h}(1;10,2,3) + \mathbf{h}(2;10,2,3) + \mathbf{h}(3;10,2,3) \end{aligned}$$

وبالتعويض عن كل حدٍّ بما يساويه نحصل على العلاقة الآتية:

$$P(X \le 3) = \sum_{i=0}^{3} \binom{5}{i} \cdot \binom{20}{10-i} \cdot \binom{25}{10}^{-1} = 0.0565 + 0.2569 + 0.3854 + 0.2372 = 0.936$$

في الواقع إنَّ إيجاد حلول مسائل تتعلَّق بالتوزيع فوق الهندسي ليس بالأمر اليسير دومًا، لا بل قد يكون بحاجة إلى تقنيات متقدمة لحلها، والمثال التالي يوضِّح لنا ذلك.

٣- في مستودع لمصنع إنتاج المصابيح الكهربائية يوجد 50000 مصباح من نوع معيّن (أي لجميع المصابيح النصيب نفسه في الاختيار) منها 500 مصباح مُعطّل، فإذا قمنا بسحب عشوائي لـ 50 مصباحاً دفعةً واحدة، وأخضعنا كل مصباح من هذه الدفعة للاختبار، فما هو احتمال ظهور خمسة مصابيح معطّلة خلال هذه الاختبارات؟

الحـل: ليكن $X$ متغيراً عشوائياً راصداً لعدد المصابيح المعطّلة التي ستظهر لدينا خلال هذه الاختبارات، فعندئذ سيكون لهذا المتغيِّر العشوائي توزيعًا فوق هندسي بمَعَالم $N = 50000$ ، $M = 500$ و $n = 50$ ، ومن ثَّم الاحتمال المطلوب هو $P(X = 5)$ حيث لدينا:

$$P(X = 5) = \mathbf{h}(5;50000,500,50) = \frac{\binom{500}{5}\cdot\binom{49500}{1995}}{\binom{50000}{50}}$$

وهنا نلاحظ الصعوبة البالغة لحساب التوافيق في كل من البسط والمقام، وحتى الحاسبات اليدوية البسيطة قد لا تستطيع تنفيذ العمليّات الحسابيّة المطلوبة للحل. لذلك سنقدِّم في نهاية هذا الفصل آليّة فعَّالة لحل مثل هذه المسائل.

### (٦ , ١ , ٨) توزيع بواسون Poisson Distribution

يتعامل هذا المتغيِّر العشوائي مع حوادث نادرة الوقوع (قدِّم أحد نماذجه كمثال في الفصل السابق)، ولذلك سنلاحظ أنَّ آليّة عمله قد ترتبط بعمل متغيِّرات عشوائية سابقة، وهو من المتغيِّرات العشوائية المهمَّة في مجال نظرية الطوابير Queuing theory (وتكوِّن أساساً لنظرية الخدمة **Service theory**) ونظرية الوثوقية Reliability theory، وقُدِّم هذا التوزيع لأول مرَّة (عام **1837**) كحالة تقاربية للتوزيع الحدَّاني من قبل الرياضياتي والفيزيائي الفرنسي بواسون (1781-1840) Siméon Denis Poisson.

#### (٦ , ١ , ٨ , ١) تعريفه

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ ، فعندئذ يُقال عن $X$ إنَّه خاضع لتوزيع بواسون بمَعْلمة $\lambda > 0$ إذا كان يأخذ قيم صحيحة غير سالبة، وقانون توزيعه مُعطى من خلال العلاقة الآتية:

$$P(X = k) = \frac{\lambda^k}{k\,!}\,\mathbf{e}^{-\lambda}\ ;\ k \in \mathbb{N}^o \qquad [6,8]$$

#### (٦ , ١ , ٨ , ٢) ملاحظات

١- سوف نستخدم الرمز $\mathbf{P}(\lambda)$ للدلالة على هذا التوزيع، وكذلك سنرمز بـ $\mathbf{p}(k;\lambda)$ للدلالة على قيمة الكتل الاحتمالية لهذا المتغير العشوائي $X$، أي إنَّ $\mathbf{p}(k;\lambda) \triangleq P(X = k)$ .

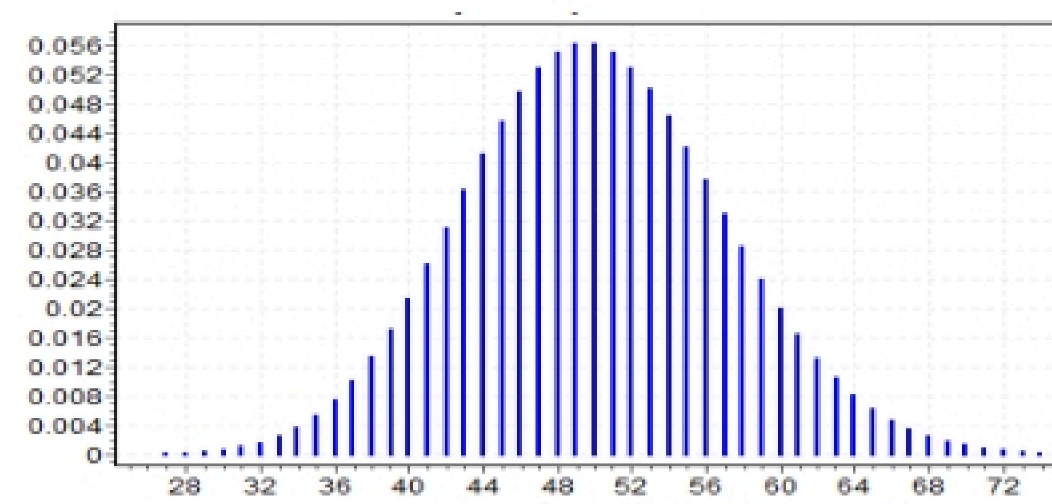
٢- يستخدم هذا النوع من التوزيعات في حساب احتمالات لحوادث نادرة الوقوع، وفي نظرية الخدمة، وكذلك في نظرية الموثوقية، وفي دراسة توزيعات الجسيمات الدقيقة أيضاً.

٣- لتوضيح العرض الجدولي لهذا المتغيِّر العشوائي سنأخذ للمَعْلمة $\lambda$ القيمة $\lambda = 1.5$ ، فيكون لجدول توزيع هذا المتغيِّر العشوائي العرض الآتي:

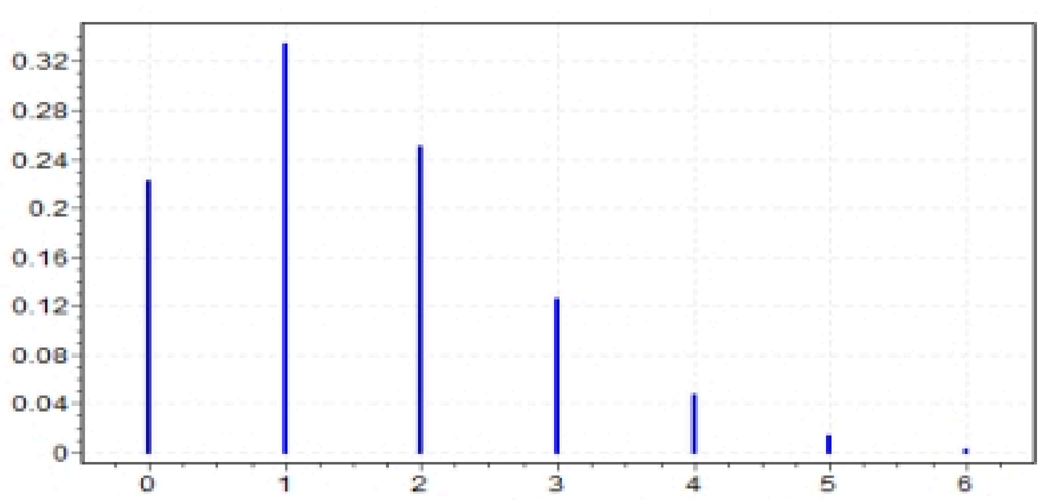
الجدول (٦,٨).

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | sum |
|---|---|---|---|---|---|---|---|
| $X$ قيم $x_i$ | 0 | 1 | 2 | 3 | 4 | 5 | ... |
| $p_i = P(X = x_i)$ | 0.223 | 0.335 | 0.251 | 0.126 | 0.047 | 0.014 | $\approx 1$ |

وأمّا التمثيل البياني لهذا المتغيِّر العشوائي من أجل قيم مختلفة لـ $\lambda$ يقدّمها الشكلين الآتيين:



الشكل (٦,١٥.أ) من أجل $\lambda = 50$



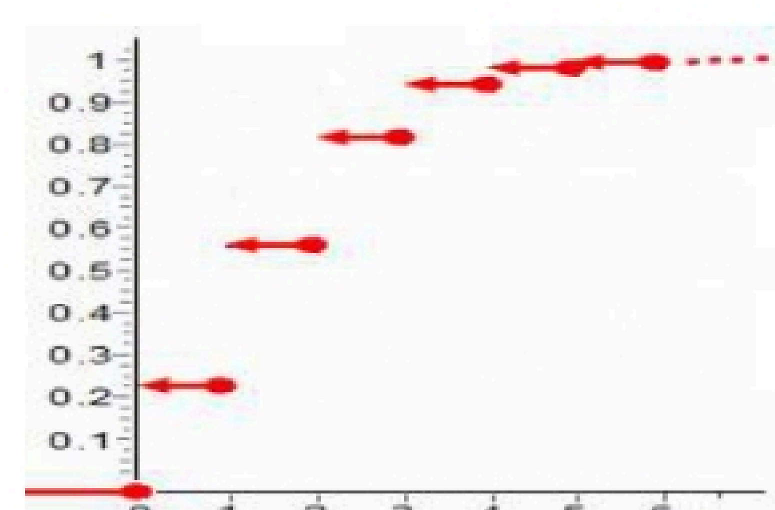
الشكل (٦,١٥.ب) من أجل $\lambda = 1.5$

٤- إنَّ دالَّة توزيع هذا المتغيِّر العشوائي نجدها باستخدام تعريف دالَّة توزيع متغيِّر عشوائي متقطِّع، حيث لدينا:

$$F_X(x) = \sum_{k=0\,;\,k<x}^{\infty} \frac{\lambda^k}{k\,!}\,\mathrm{e}^{-\lambda} \qquad ; x \in \mathbb{R}$$

ومن أجل $\lambda = 1.5$ يكون لها العرض الآتي:

$$F_X(x) = \begin{cases} 0 & \text{for } x \le 1 \\ 0.2231 & \text{for } 1 < x \le 2 \\ 0.5577 & \text{for } 2 < x \le 3 \\ 0.8088 & \text{for } 3 < x \le 4 \\ 0.9344 & \text{for } 4 < x \le 5 \\ \vdots & \vdots \quad \vdots \quad \vdots \quad \vdots \\ 1 & \text{for } x \to \infty \end{cases}$$



الشكل (٦,١٦) دالَّة التوزيع الاحتمالية من أجل $\lambda = 1.5$ .

يُلاحظ هنا الصعوبة في حساب الدالَّة $F_X$ من أجل القيم الكبيرة للوسيط $\lambda$، ولكن يمكن حسابها باستخدام ما يسمَّى بـ **تكامل غاما الناقص** والذي له الصيغة $\Gamma_k = \dfrac{1}{k\,!}\displaystyle\int_{\lambda}^{+\infty} \mathrm{e}^{-t}\, t^k\, dt$ حيث يبرهن على صحة العلاقة الآتية:

$$F_X(x) = \frac{1}{\lfloor x \rfloor !}\int_{\lambda}^{+\infty} \mathrm{e}^{-t}\, t^{\lfloor x \rfloor}\, dt$$

### (٦,١,٨,٣) أمثلة

١- إذا علمت أنَّ عدد الأهداف التي يمكن لمنتخب كرة قدم أن يسجلّها في مباراة واحدة يخضع لتوزيع بواسون بمَعْلمة $\lambda = 1$ (الوحدة لـ $\lambda$ هنا هي الهدف)، فعندئذ:

أ- ما هو احتمال أن يسجِّل هذا الفريق أربعة أهداف في مباراة مقبلة؟

ب- ما هو احتمال أن يسجِّل هذا الفريق هدفًا واحدًا على الأقل في مباراة مقبلة؟

**الحـل**: بفرض أنَّ $X$ متغيِّرًا عشوائيًّا راصدًا لعدد الأهداف التي سيسجلها الفريق في مباراة مقبلة، فعندئذ سيكون $X$ خاضعاً لتوزيع بواسون بمَعْلمة $\lambda = 1$ ، ومن ثمَّ يكون لدينا:

أ- احتمال أن يسجِّل هذا الفريق أربعة أهداف في مباراة مقبلة هو:

$$P(X = 4) = \mathbf{p}(4;1) = \frac{1^4}{4!}\mathbf{e}^{-1} = 0.01533$$

ب- احتمال أن يسجِّل هذا الفريق هدف واحد على الأقل في مباراة مقبلة هو:

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 1 - \mathbf{p}(0;1) = 1 - 0.369 = 0.631$$

٢- إذا علمت أنَّ عدد الزوار خلال دقيقة واحدة إلى مخدِّم الويب يتبع توزيع بواسون بمَعْلمة $\lambda = 7$ (الوحدة لـ $\lambda$ هنا هي زائر)، فعندئذ:

أ- ما هو احتمال وجود ثلاثة زوار على الأكثر خلال فترة دقيقة ما من زمن عمل المخدِّم؟

ب- ما هو احتمال أن تمرّ دقيقة من زمن عمل المخدِّم دون دخول أي زائر عليه؟

**الحـل**: بفرض أنَّ $X$ متغيِّرًا عشوائيًّا راصدًا لعدد الزوار الداخلين إلى خدمة الويب في كل دقيقة، فعندئذ سيكون $X$ خاضعاً لتوزيع بواسون بمَعْلمة $\lambda = 7$ ، ومن ثمَّ يكون لدينا من أجل الطلب:

أ- الاحتمال المطلوب هو:

$$\begin{aligned} P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\ &= \mathbf{p}(0;7) + \mathbf{p}(1;7) + \mathbf{p}(2;7) + \mathbf{p}(3;7) \\ &= 0.0009 + 0.0064 + 0.0223 + 0.0521 = 0.0817 \end{aligned}$$

ب- الاحتمال المطلوب هو:

$$P(X = 0) = \mathbf{p}(0;7) = 0.0009$$

٣- عند تقاطع مُعيَّن لطريقين رئيسيين وجد أنَّ عدد الحوادث المرورية التي تحصل عنده أسبوعياً تتوزع بواسونياً بمَعْلمة $\lambda = 0.5$ ، ولنقم باستخدام التوزيع البواسوني لحساب:

أ- احتمال وقوع خمسة حوادث عند هذا التقاطع خلال أسبوع ما؟

ب- احتمال وقوع حادثين على الأقل عند هذا التقاطع خلال أسبوع ما؟

**الحـل**: من أجل الطلب:

أ- إنَّ الاحتمال المطلوب هو:

$$P(X = 5) = \mathbf{p}(5;0.5) = 0.000158$$

ب- إنَّ الاحتمال المطلوب هو الآتي:

$$\begin{aligned} P(X \geq 2) &= 1 - P(X < 2) = 1 - [P(X = 0) + P(X = 1)] \\ &= 1 - [\mathbf{p}(0;0.5) + \mathbf{p}(8;0.5)] = 1 - [0.6065 + 0.3033] = 0.0902 \end{aligned}$$

تطبيقات أخرى على استخدام التوزيع البواسوني نجدها بعد فقرة تقارب التوزيعات الاحتمالية في نهاية هذا الفصل.

سنكتفي بهذا القدر من المعلومات حول بعض المتغيرات العشوائية المتقطِّعة حيث يوجد الكثير منها، ومن يودّ الاطلاع على المزيد يمكنه الرجوع إلى بعض المراجع المذكورة في نهاية هذا الكتاب.

## (٦,٢) توزيعات احتمالية مستمرة

**Continuous Probability Distributions**

فيما يلي سنقدِّم بعض التوزيعات الاحتمالية المستمرّة الشهيرة التي سيكون لها دور مهمّ في مجال العشوائيات، وسنبدأ هذه التوزيعات بالتوزيع الآتي.

### (٦,٢,١) التوزيع المنتظم المستمر Continuous Uniform Distribution:

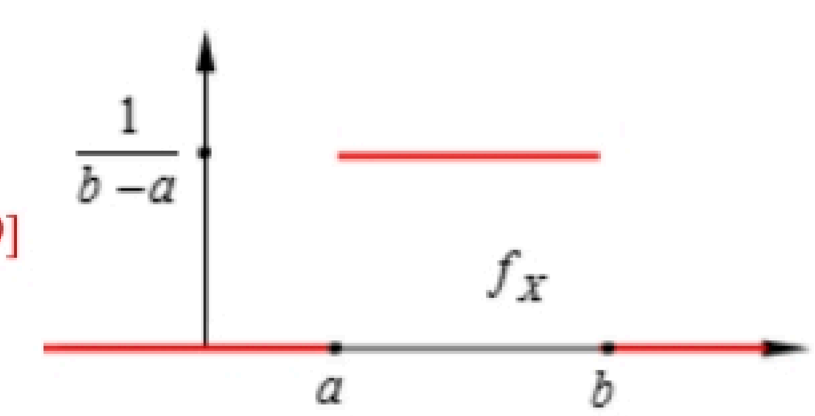
يعدّ هذا التوزيع من أهمّ وأكثر المتغيِّرات العشوائية المستخدمة في توليد الأرقام العشوائية.

#### (٦,٢,١,١) تعريفه

ليكن $X$ متغيِّراً عشوائياً مستمراً فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$ ، فعندئذٍ يُقال عن $X$ إنَّه خاضع للتوزيع **المنتظم المستمر** على فترة محدودة $[a\,,b] \subset \mathbb{R}$ إذا كانت دالَّة كثافته الاحتمالية $f_X$ مُعطاة من خلال العلاقة الآتية من أجل كل $x \in \mathbb{R}$ :

$$f_X(x) = \begin{cases} 0 & for \quad x < a \\ \dfrac{1}{b-a} & for \quad a \le x \le b \\ 0 & for \quad x > b \end{cases}$$

[6,9]




الشكل (٦,١٧)

ولرسمها البياني الشكل السابق حيث نلاحظ أنَّ الدالَّة $f_X$ تعاني من نقطتي انقطاع (من النوع الأول) في كلٍ من النقطتين $a$ و $b$، ومن ثَمَّ ستكون دالَّة توزيع هذا المتغيِّر العشوائي غير قابلة للمفاضلة في كلٍ من النقطتين $a$ و $b$.

#### (٦,٢,١,٢) ملاحظات

١- سوف نستخدم الرمز $\mathbf{U}([a,b])$ للدلالة على هذا التوزيع.

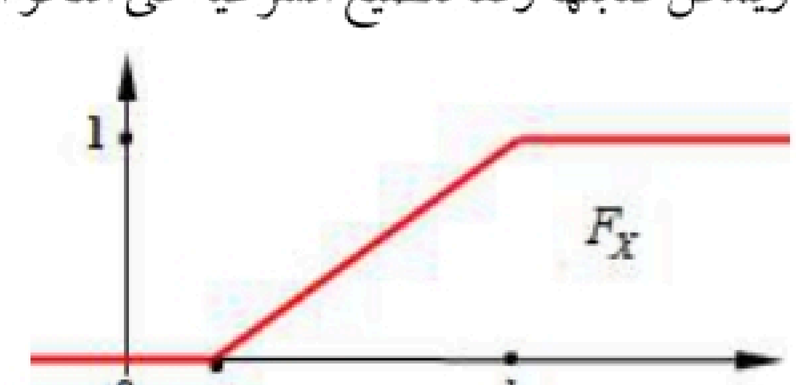
٢- من أهمّ استخدامات هذا المتغيِّر العشوائي توليد الأرقام العشوائية.

٣- باستخدام تعريف دالَّة التوزيع الاحتمالية لمتغيِّر عشوائي مستمر يكون لدالَّة توزيع هذا المتغيِّر العشوائي العرض الآتي من أجل كل $x \in \mathbb{R}$ :

$$F_X(x) = \frac{1}{b-a}\int_{-\infty}^{x} \mathbf{I}_{[a,b]}(t)\ dt \ = \frac{x-a}{b-a}\,\mathbf{I}_{(a,b]}(x) + \mathbf{I}_{(b,\infty)}(x)$$

ويمكن كتابتها وفقاً للصيغ الشرطية على النحو الآتي أيضاً:

$$F_X(x) = \begin{cases} 0 & for \quad x \le a \\ \dfrac{x-a}{b-a} & for \quad a < x \le b \\ 1 & for \quad x > b \end{cases}$$




الشكل (٦,١٨) الرسم البياني لدالَّة التوزيع الاحتمالية.

**(٣, ١, ٢, ٦) أمثلة**

١ – لنفترض أن المدّة الزمنية التي تحتاجها إقامة الأفراح أو الأتراح في استرحة ما تخضع للتوزيع المنتظم على الفترة $[0,5]$ (مقدَّرة بالساعات)، فعندئذ لنحسب احتمال أن تُشغل هذه الاستراحة لمدَّة لا تتجاوز أربع ساعات.

الحل: من أجل ذلك ليكن $X$ متغيّراً عشوائياً راصداً لزمن انشغال الاستراحة، فعندئذ يكون لدالَّة الكثافة الاحتمالية لتوزيع زمن انشغال الاستراحة العرض الآتي:

$$f_X(x) = \begin{cases} \dfrac{1}{5} & for\ \ 0 \le x \le 5 \\ 0 & otherwise \end{cases}$$

ومن ثمَّ يكون الاحتمال المطلوب هو:

$$P(X \le 4) = \int_0^4 \frac{1}{5}\,dx = \frac{1}{5}\int_0^4 dx = \frac{4}{5} = 0.8$$

٢– في مصنع لإنتاج السجاد يوجد آلة لنسج السجاد بحيث يخضع زمن إنتاجها (مقدَّرة بالدقيقة) للتوزيع المنتظم على الفترة $[30,45]$، فإذا أخذنا سجادة ما من المنتج لهذا المصنع فما هو احتمال أن يكون الزمن الذي استغرقه إنتاج هذه السجادة لا يتجاوز 35 دقيقة؟

الحل: من أجل ذلك ليكن $X$ متغيّراً عشوائياً راصداً للمدَّة الزمنية اللازمة لإنتاج سجادة من قبل الآلة، فعندئذ يكون لدالَّة الكثافة الاحتمالية لتوزيع زمن الإنتاج العرض الآتي:

$$f_X(x) = \begin{cases} \dfrac{1}{15} & for\ \ 30 \le x \le 45 \\ 0 & otherwise \end{cases}$$

وأمّا الاحتمال المطلوب فهو $P(X \le 35)$، وقيمته تساوي:

$$P(X \le 35) = \int_{30}^{35} \frac{1}{15}\,dx = \frac{1}{15}\int_{30}^{35} dx = \frac{5}{15} = 0.\overline{3}$$

## (٢, ٢, ٦) التوزيع الأسي Exponential Distribution

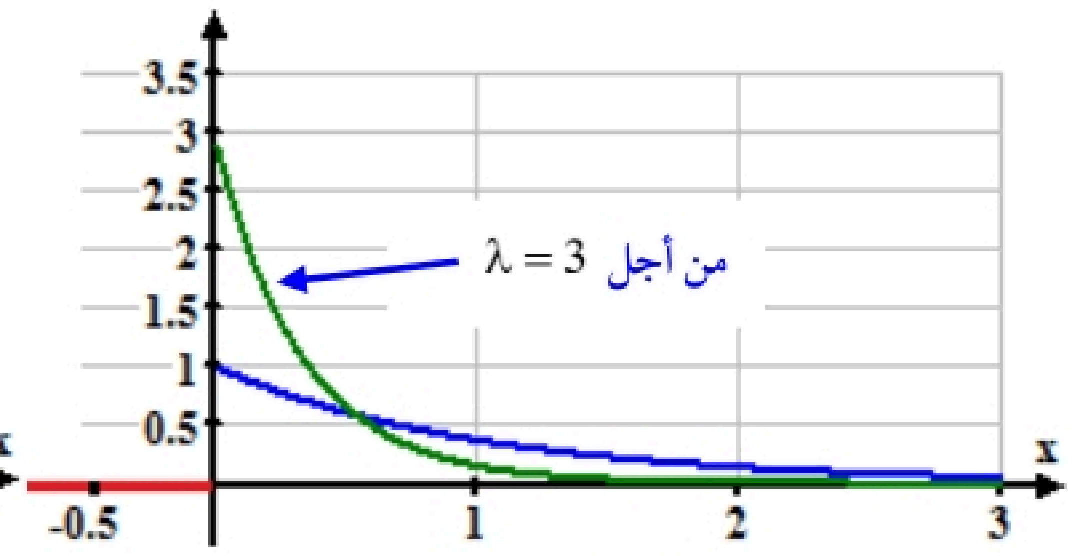
يعدّ هذا المتغيِّر العشوائي من المتغيِّرات العشوائية المهمَّة جداً في مجال دراسة نظرية الوثوقية ونظرية الطوابير، ويستخدم في مجال توليد الأرقام العشوائية أيضاً.

**(١, ٢, ٢, ٦) تعريفه**

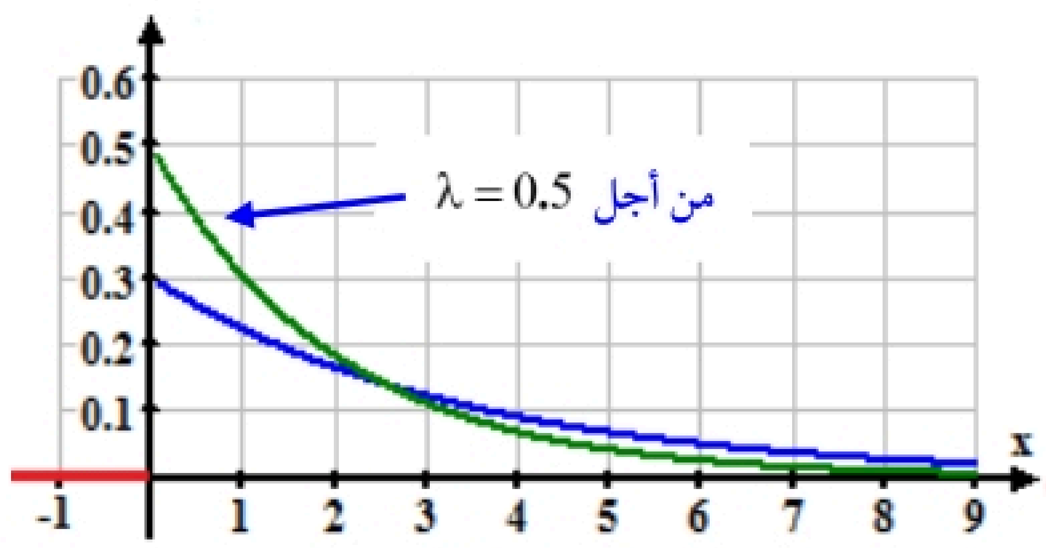
ليكن $X$ متغيِّراً عشوائياً مستمراً فوق فضاء احتمالي $[\Omega, \mathcal{A}, P]$، فعندئذٍ يُقال عن $X$ إنَّه خاضع للتوزيع **الأسي** بمَعْلمة $0 < \lambda$ إذا كانت دالَّة كثافته الاحتمالية $f_X$ مُعطاة من خلال العلاقة الآتية:

$$f_X(x) = \lambda \mathbf{e}^{-\lambda x} \cdot \mathbf{I}_{[0,+\infty)}(x) \qquad ; \forall x \in \mathbb{R} \qquad [6,10]$$

والرسم البياني لدالَّة الكثافة الاحتمالية لهذا المتغيِّر العشوائي يقدّمها الشكلين الآتيين من أجل قيمتين مختلفتين لـــ $\lambda$ .




الشكل (١٩, ٦.أ) من أجل $\lambda = 3, \lambda = 1$




الشكل (١٩, ٦.ب) من أجل $\lambda = 0.5, \lambda = 0.3$

فنلاحظ أنَّ الدالَّة $f_X$ تعاني من نقطة انقطاع (من النوع الأول) في المبدأ، ومن ثمَّ ستكون دالَّة توزيع هذا المتغيِّر العشوائي غير قابلة للمفاضلة في المبدأ.

### (٢, ٢, ٢, ٦) ملاحظات

١- سوف نستخدم الرمز $\mathbf{Ex}(\lambda)$ للدلالة على هذا التوزيع.

٢- يستخدم هذا النوع من المتغيِّرات العشوائية في نظرية الوثوقية ونظرية الخدمة، وكذلك في توليد الأرقام العشوائية.

٣- من الخصائص المهمَّة لهذا التوزيع أنَّه لا يمتلك ذاكرة (كما هو الحال لدى التوزيع الهندسي)، ويعبَّر عن ذلك بالعلاقة الآتية:

$$P\left(X \geq x+t \mid X \geq t\right)=P\left(X \geq x\right) \qquad ; x, t>0$$ [6,11]

ولإثباتها سنستخدام مفهوم الاحتمال الشرطي حيث يمكننا أن نكتب الآتي.

بسبب أنَّ $\{\omega \in \Omega ; X(\omega)>x+t\} \supseteq \{\omega \in \Omega ; X(\omega)>t\}$ فإنَّه سيكون لدينا:

$$P\left(X \geq x+t \mid X \geq t\right)=\frac{P\left(X \geq x+t, X \geq t\right)}{P\left(X \geq t\right)}=\frac{P\left(X \geq x+t\right)}{P\left(X \geq t\right)}=\frac{1-P\left(X<x+t\right)}{1-P\left(X<t\right)}$$

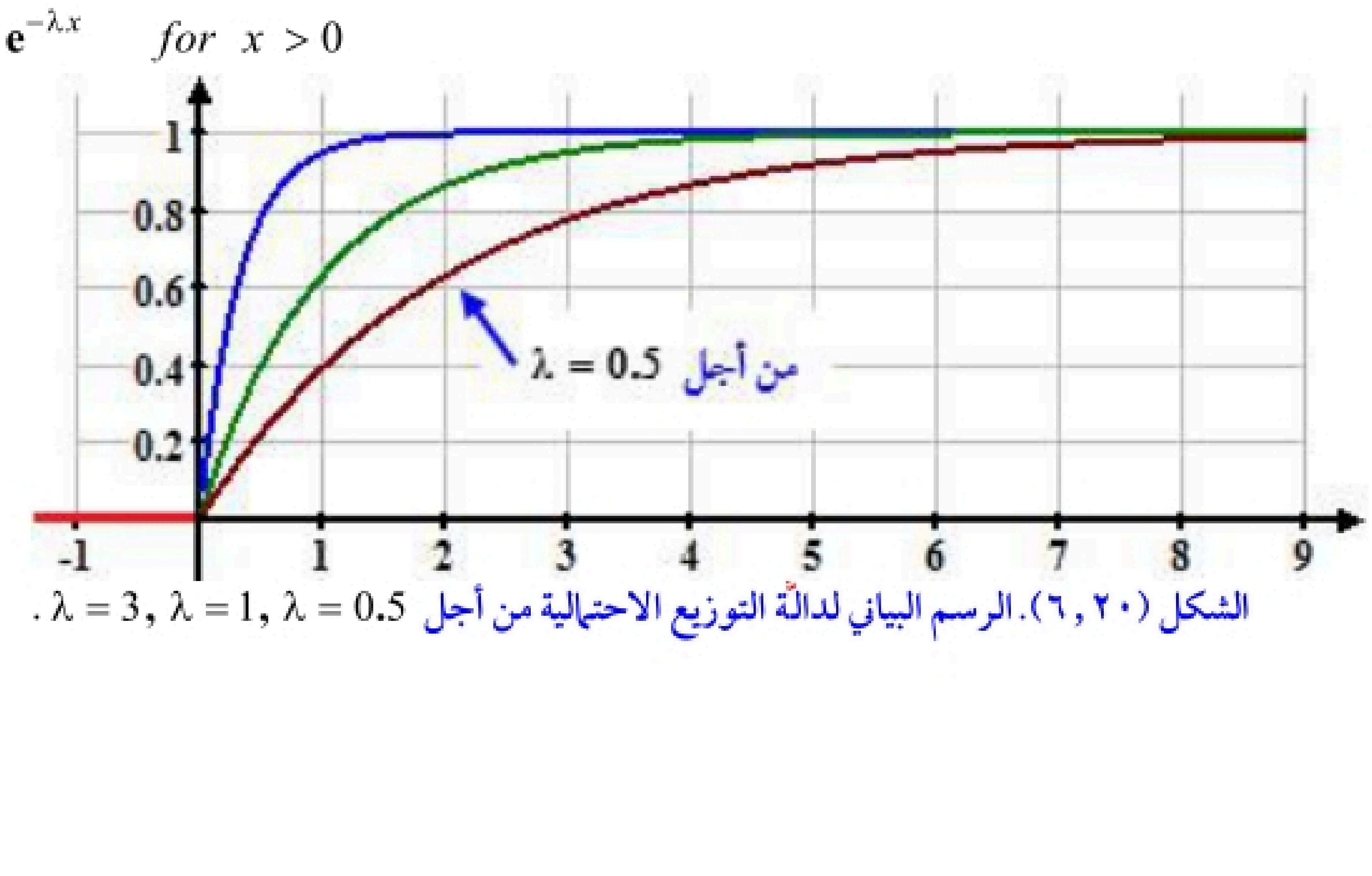
$$=\frac{\mathbf{e}^{-\lambda(x+t)}}{\mathbf{e}^{-\lambda t}}=\mathbf{e}^{-\lambda x}=P\left(X \geq x\right)$$

٤- باستخدام تعريف دالَّة التوزيع الاحتمالية لمتغيِّر عشوائي مستمر يكون لدالَّة توزيع هذا المتغيِّر العشوائي العرض الآتي:

$$F_X(x)=\int_{-\infty}^{x} \lambda \mathbf{e}^{-\lambda t} \mathbf{I}_{[0,+\infty)}(t)\, dt=\int_{0}^{x} \lambda \mathbf{e}^{-\lambda t}\, dt=\left(1-\mathbf{e}^{-\lambda x}\right) \mathbf{I}_{[0,+\infty)}(x) \qquad ; \forall x \in \mathbb{R}$$

والتي يمكن كتابتها من أجل كل $x \in \mathbb{R}$ على النحو الآتي أيضاً:

$$F_X(x)=\begin{cases}0 & \text{for } x \leq 0 \\ 1-\mathbf{e}^{-\lambda x} & \text{for } x>0\end{cases}$$




الشكل (٢٠, ٦). الرسم البياني لدالَّة التوزيع الاحتمالية من أجل $\lambda = 0.5, \lambda = 1, \lambda = 3$.

**(٣، ٢، ٢، ٦) أمثلة**

١- إذا علمت أنَّ الوقت اللازم لإصلاح سيّارة في ورشة صيانة هو متغيِّر عشوائي أسّي بمَعْلمة $\lambda = 1.5$ ساعة، فعندئذ:

أ- لنحسب احتمال ألاّ يستغرق وقت إصلاح سيّارة في هذه الورشة أكثر من ثلاث ساعات.

ب- لنحسب احتمال أن يستغرق إصلاح سيّارة في هذه الورشة أكثر من أربع ساعات.

ج- ما هو احتمال أن يستغرق وقت إصلاح السيّارة أكثر من أربع ساعات علماً أنَّ عامل الإصلاح بدأ عمله في إصلاح السيارة منذ ثلاث ساعات؟

**الحـل**: من أجل الطلب ( أ ) لدينا الاحتمال المطلوب هو:

$$\begin{aligned} P(X \le 3) &= P(X < 3) = F_X(3) \\ &= \int_{-\infty}^{3} \lambda\, \mathbf{e}^{-\lambda t}\, \mathbf{I}_{[0,+\infty)}(t)\; dt = \int_{0}^{3} 1.5\, \mathbf{e}^{-1.5t}\, dt \\ &= 1 - \mathbf{e}^{-1.5(3)} = 0.98889 \end{aligned}$$

ب- من أجل الطلب (ب) لدينا الاحتمال المطلوب هو:

$$\begin{aligned} P(X > 4) &= 1 - P(X \le 4) = 1 - P(X < 4) = 1 - F_X(4) \\ &= 1 - \left[1 - \mathbf{e}^{-1.5(4)}\right] = 1 - 0.9975 = 0.0025 \end{aligned}$$

ج- إمن أجل الطلب (ج) لدينا الاحتمال المطلوب هو $P(X > 4 \mid X > 3)$، وبما أنَّ التوزيع الأسي لا يمتلك ذاكرة فإنَّه سيكون لدينا:

$$\begin{aligned} P\left(X > 1+3 \mid X > 3\right) &= P\left(X > 1\right) \\ &= 1 - F_X(1) = 1 - 0.777 = 0.223 \end{aligned}$$

٢- لنفترض أن أجهزة تحتوي على نوع معين من العناصر الإلكترونية التي يمكن أن تتعطّل خلال فترة زمنية $T$ (مقدرةً بالسنوات). فإذا علمنا أنَّ التوزيع الاحتمالي للمتغير العشوائي $T$ هو توزيع أسي بمَعْلمة $\lambda = 1$ سنة، وأنّه قد تمّ تثبيت خمسة عناصر في أجهزة مختلفة وبشكل مستقلّ بعضها عن البعض الآخر، فما هو احتمال أن يبقى على الأقل عنصرين يعملان حتى نهاية السنوات الخمس القادمة؟

**الحـل**: من أجل ذلك سناخذ $X$ متغيّراً عشوائياً راصداً لعدد الأجهزة التي ستبقى تعمل حتى نهاية السنوات الخمس القادمة، فعندئذ سيكون لهذا المتغيّر العشوائي توزيعًا حدّانيًا بمَعْلمتين $n = 5$ و $p = ?$ يجب حسابه من التوزيع الاحتمالي للمتغيّر العشوائي $T$، وأمّا الاحتمال المطلوب فهو $P(X \ge 2)$. حيث لدينا:

$$\begin{aligned} p = P(T > 5) &= 1 - P(T \le 5) = 1 - P(T < 5) \\ &= 1 - F_T(5) = 1 - \left(1 - \mathbf{e}^{-5}\right) = 1 - 0.993 \approx 0.007 \end{aligned}$$

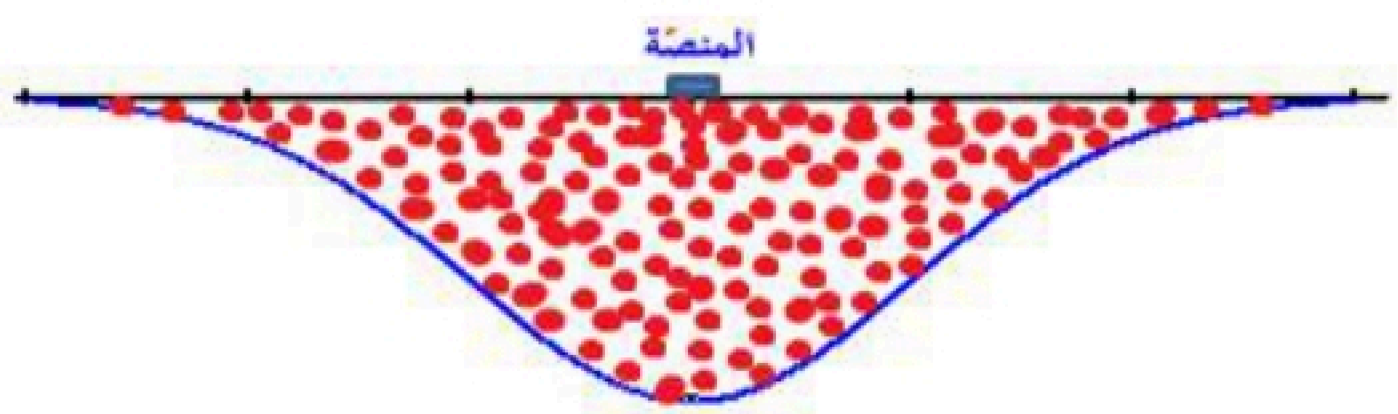
ومنه يكون الاحتمال المطلوب هو:

$$\begin{aligned} P(X \ge 2) &= 1 - P(X < 2) = 1 - \left[P(X = 0) + P(X = 1)\right] \\ &= 1 - \left[\mathbf{b}(0; 0.007) + \mathbf{b}(1; 0.007)\right] \\ &= 1 - \left[0.965 + 0.033\right] = 1 - 0.998 = 0.002 \end{aligned}$$

### (٦,٢,٣) التوزيع الطبيعي Normal Distribution

إنَّ التوزيع الطبيعي (أو ما يُعرف باسم التوزيع الغاوسي Gaussian Distribution أيضاً) يُعدّ من أكثر التوزيعات الاحتمالية أهمية وفائدة في حقل الإحصاء بأكمله، وهذا التوزيع يصف العديد من الظواهر التي تحدث في الطبيعية، والصناعة، والبحوث العلميّة على حدّ سواء، فعلى سبيل المثال:

- وجد أنَّ الكثير من الظواهر في الطبيعية تتوزع احتمالياً وفقاً لهذا التوزيع، فمنها ظاهرة التجمهر أمام موضع مُحدَّد في منطقة مَفتوحة، حيث يتجمهّر النّاس أمام منصَّة في منطقة مفتوحة، ومن الطبيعي أن يفضِّل معظَّمهم الجلوس (أو الوقوف) قريباً من المنصَّة بقدر المستطاع، فيكون لهم التوضّع الموضَّح في الرسم الآتي (في معظّم الحالات).




ولهذا السبب دعي التوزع المقدَّم آنفاً باسم "التوزيع الطبيعي"، وذلك لأنَّه نتج عن سلوك فطري لدى النّاس.

- القياسات الفيزيائية في مجالات عديدة كقياسات الأجزاء الصناعية يكون لتوزيعها الاحتمالي (وفي كثير من الأحيان) تقريب جيد مع التوزيع الطبيعي.
- الأخطاء المرتكبة لدى أخذ القياسات المعمليّة يمكن تقريبها بشكل جيد من التوزيع الطبيعي، وهناك أمثلة كثيرة أخرى لا مجال لذكرها في هذه العُجالة من التقديم.

لقد كان الرياضياتي الفرنسي موافير Abraham de Moivre (1667-1754) أول من وضع المعادلة الرياضياتية للمنحنى الطبيعي في عام 1733، ولكن جاء بعده الرياضياتي الألماني غاوس Johann Carl Friedrich Gauss (1777-1855) ليستنتج المعادلة الرياضياتية للمنحنى الطبيعي من دراسة الأخطاء في القياسات المتكرِّرة للكميّة نفسها، ولهذا أُطلق على هذا التوزيع اسم توزيع غاوس تكريماً له، وبسبب الأهمية الكبيرة لهذا التوزيع سوف نستفيض قليلاً في تقديمه.

#### (٦,٢,٣,١) تعريفه

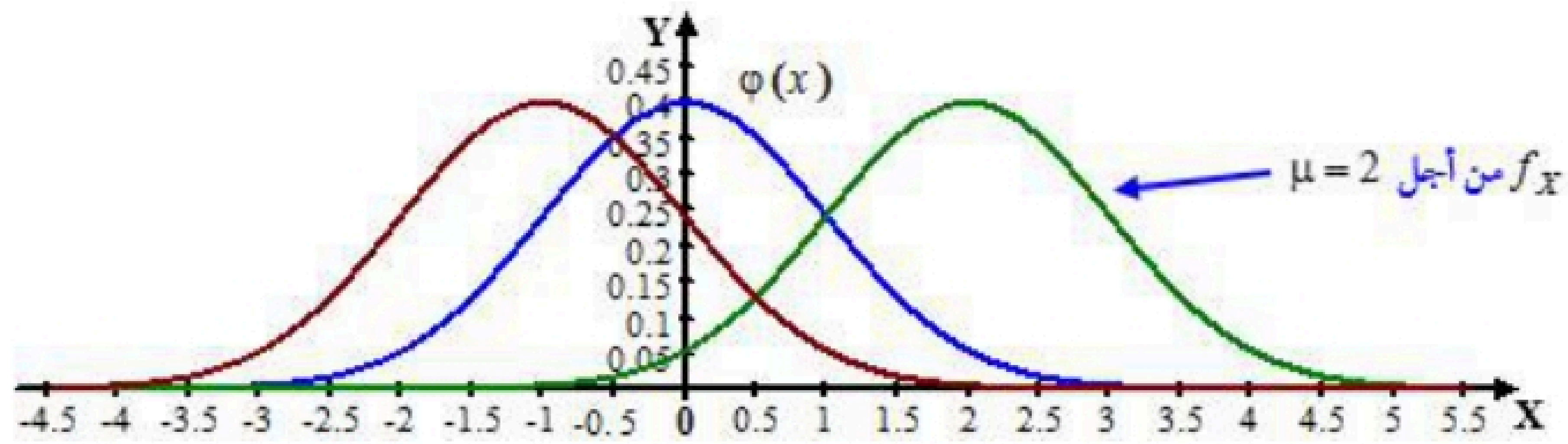
ليكن $X$ متغيِّراً عشوائياً مستمراً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ ، فعندئذ يُقال عن $X$ إنَّه خاضع للتوزيع الطبيعي بمَعْلمتين $\mu \in \mathbb{R}$ و $\sigma > 0$ إذا كانت دالَّة كثافته الاحتمالية $f_X$ مُعطاة من خلال العلاقة الآتية:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \mathbf{exp}\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad ;\forall x \in \mathbb{R} \qquad [6,12\text{-}a]$$

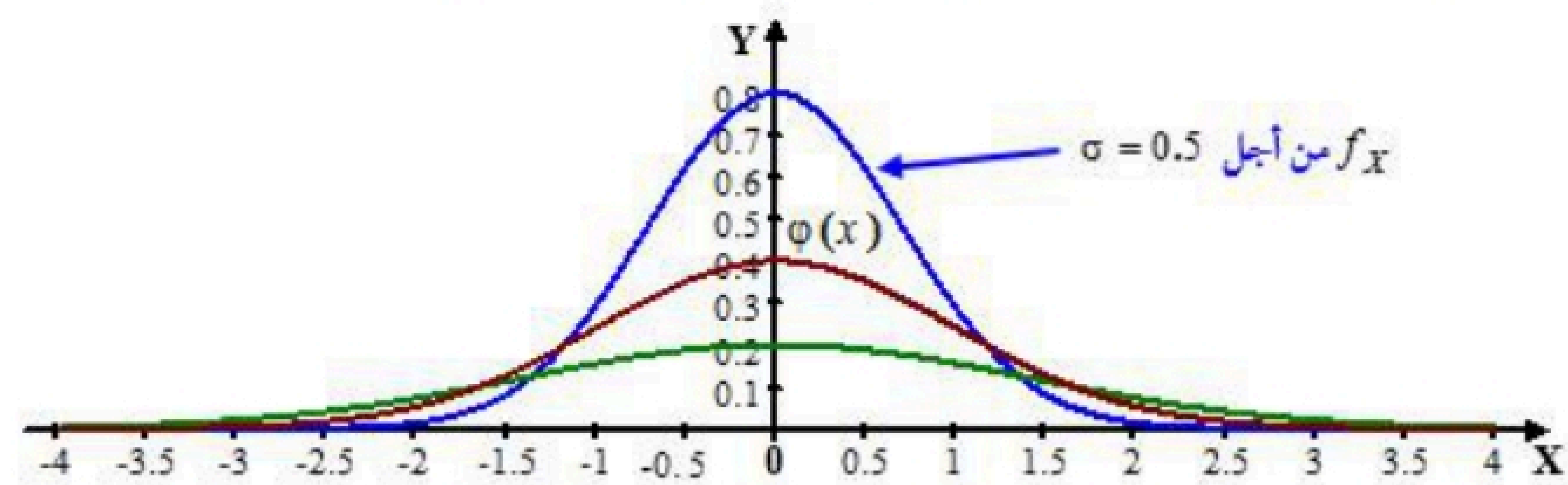
وفي الحالة الخاصة عندما تكون $\mu = 0$ و $\sigma = 1$ فإنَّه يصبح للعلاقة السابقة العرض الآتي:

$$f_X(x) = \frac{1}{\sqrt{2\pi}} \mathbf{e}^{-\frac{x^2}{2}} \qquad ;\forall x \in \mathbb{R} \qquad [6,12\text{-}b]$$

ويُقال حينئذ إنَّ المتغيِّر العشوائي $X$ خاضع للتوزيع الطبيعي المعّياري Standard Normal Distribution ويُرمز لدالَّة الكثافته الاحتمالية الموافقة له بـ $\varphi(x)$ ، ويصبح منحنى هذه الدالَّة متناظراً بالنسبة إلى المحور $\mathbf{oY}$ كما هو واضح في الشكلين الآتيين الذين يقدِّمان لنا الرسم البياني لهما من أجل قيم مختلفة لـ $\mu \in \mathbb{R}$ و $\sigma > 0$ :




الشكل (٦,٢١ . أ) دالّة كثافته الاحتمالية للتوزيع الطبيعي من أجل $\sigma = 1$ للجميع و $\mu = 2$ ، $\mu = 0$ و $\mu = -1$




الشكل (٦,٢١ . ب) دالّة كثافته الاحتمالية للتوزيع الطبيعي من أجل $\mu = 0$ للجميع و $\sigma = 0.5$ ، $\sigma = 1$ و $\sigma = 2$

فنلاحظ هنا أنَّ الدالَّة $f_X$ مستمرّة على $\mathbb{R}$ ، ولذلك ستكون دالَّة التوزيع الاحتمالية لهذا التوزيع مستمرّة وقابلة للمفاضلة عند كل نقطة $x$ من $\mathbb{R}$ . كما نلاحظ أنَّ تغيّر قيمة $\mu$ توافق الذروة لمنحنى دالَّة الكثافة الاحتمالية، ويؤدي إلى انسحاب منحنى دالَّة الكثافة الاحتمالية لهذا التوزيع يُمنةً أو يُسرة وذلك بحسب القيمة التي تأخذها هذه المَعْلمة، وأنَّ تغيّر قيمة $\sigma$ يؤدي إلى تدبب أو انبساط منحنى دالَّة الكثافة الاحتمالية لهذا التوزيع وذلك بحسب القيمة التي تأخذها هذه المَعْلمة، وهنا نشير إلى أنَّه يُقال عن منحنى دالَّة كثافته الاحتمالية للتوزيع الطبيعي إنَّه **منبسط** إذا كان $\sigma > 1$، وفي حال كان $\sigma < 1$ فعندئذ يُقال عن هذا المنحنى إنَّه **مدبَّب**، وأمّا إذا كان $\sigma = 1$ فعندئذ يُقال عن هذا المنحنى إنَّه **مُعتدل التفلطح**. لذلك ينظر في هذه الحالة إلى شكل منحنى دالَّة كثافته الاحتمالية للتوزيع الطبيعي على أنَّه مقياس الاعتدال لأشكال التوزيعات (من حيث التدبب والانبساط) كما سبق وذكرنا ذلك في الفصل الثاني.

### (٦,٢,٣,٢) ملاحظات

١- يُرمز للتوزيع الطبيعي ذي المَعْلمتين $\mu$ و $\sigma$ بالرمز $\mathbf{N}(\mu,\sigma)$ (وبعض المراجع تستخدم الرمز $\mathbf{N}(\mu,\sigma^2)$ للدلالة عليه).

٢- عندما يكون المتغيِّر العشوائي $X$ خاضع للتوزيـع الطبيعي المعّياري فإنَّه يُرمز لدالَّة توزيعه بالرمز $\Phi(x)$ عوضاً عن $F_X(x)$ (ولو أنَّ هذا الرمز الأخير يستخدم بين الحين والآخر).

٣- يستخدم هذا النوع من المتغيِّرات العشوائية في قطاعات علميّة كثيرة جداً منها على سبيل المثال لا الحصر: توزيع حساب الأخطاء، الاختبارات الإحصائية، توليد الأرقام العشوائية و...

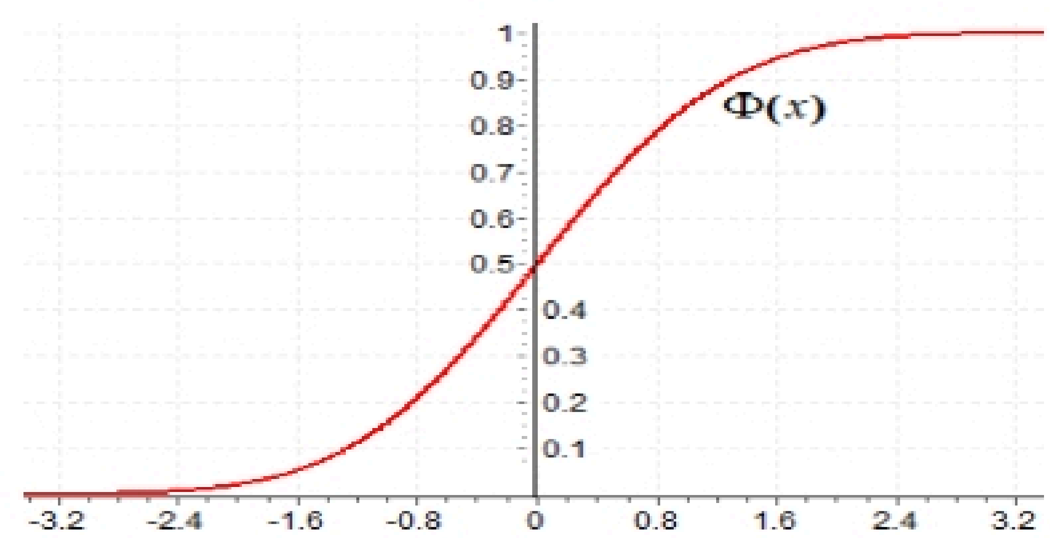
٤- إنَّ لدالَّة التوزيع الطبيعي ذي المَعْلمتين $\mu$ و $\sigma$ العرض الآتي:

$$F_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} \mathbf{e}^{-\frac{(t-\mu)^2}{2\sigma^2}}\, dt \qquad ;\forall x \in \mathbb{R}$$

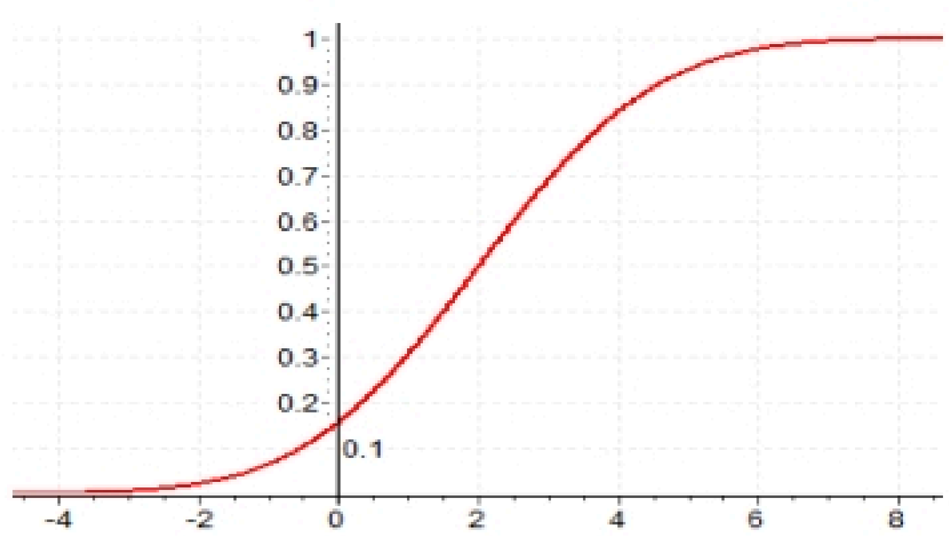
وأمّا من أجل الحالة الخاصة عندما يكون المتغيِّر العشوائي $X$ خاضعًا للتوزيع الطبيعي المعياري فإنَّه يصبح لدينا:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \mathbf{e}^{-\frac{t^2}{2}}\, dt \qquad ;\forall x \in \mathbb{R}$$

ومن أجل قيمتين مختلفتين لـ $\mu$ و $\sigma$ نجد لهذه الدالّة الرسمين البيانيين الآتيين:




الشكل (٦,٢٢ .أ) دالّة اللتوزيع الطبيعي المعياري



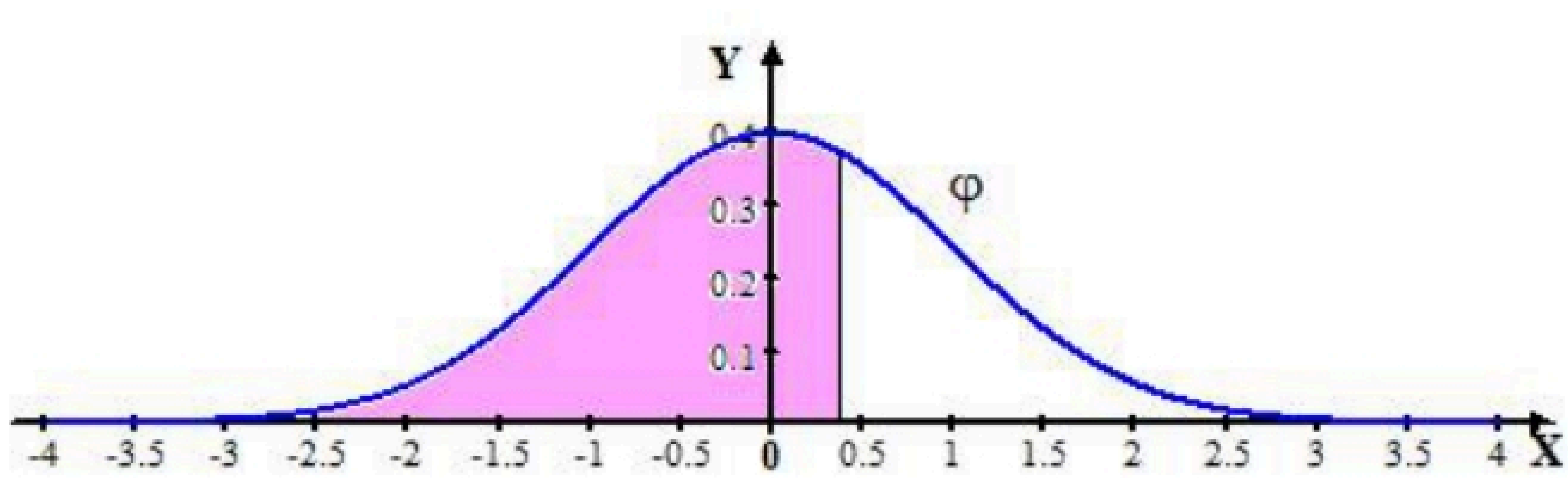
الشكل (٦,٢٢ . ب) دالّة اللتوزيع الطبيعي من أجل $\mu = 2$ للجميع و $\sigma = 2$ .

نشير هنا إلى أنّه يوجد في آخر هذا الكتاب جدولان لقيم هذه الدالّة، أحدهما من أجل القيم الموجبة لـ $z$ مبتدأة من الصفر وبتزايد يساوي 0.01، والآخر من أجل القيم السالبة لـ $z$ مبتدأة من الصفر وبتناقص يساوي 0.01، وهذه القيم تُمكِّننا من حساب احتمالات متعلّقة بمتغيِّر عشوائي طبيعي معّياري، فعلى سبيل المثال لو أخذنا $Z$ متغيِّراً عشوائياً بحيث $Z \sim \mathbf{N}(0,1)$ ، فعندئذ يكون لدينا $P(Z < 0.35) = \Phi(0.35) = 0.6368$ ، ونجدها من جدول القيم الموجبة لهذا التوزيع كما يوضِّحها العرض الآتي:

الجدول (٦,٩)

| Z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |

وتمثِّلها المساحة المظللة في الشكل الآتي:




الشكل (٦,٢٣)

سنرمز بـ $z_\alpha$ للقيمة على المحور الأفقي **oX** التي يقع على يسارها وتحت منحنى دالّة الكثافة الاحتمالية للتوزيع الطبيعي

المعياري مساحة قدرها α، فعلى سبيل المثال لدينا في الشكل السابق:

$$z_{0.6368} = 0.35$$

كما يمكن لهذه القيم التي في الجدول أن تساعدنا في تعيين القيمة $z$ لمتغيِّر عشوائي طبيعي معياري $Z$ إذا كانت قيمة الاحتمال $P(Z < z)$ معلومة، فعلى سبيل المثال لو أخذنا $Z$ متغيِّراً عشوائياً بحيث $Z \sim \mathbf{N(0,1)}$، وكان لدينا:

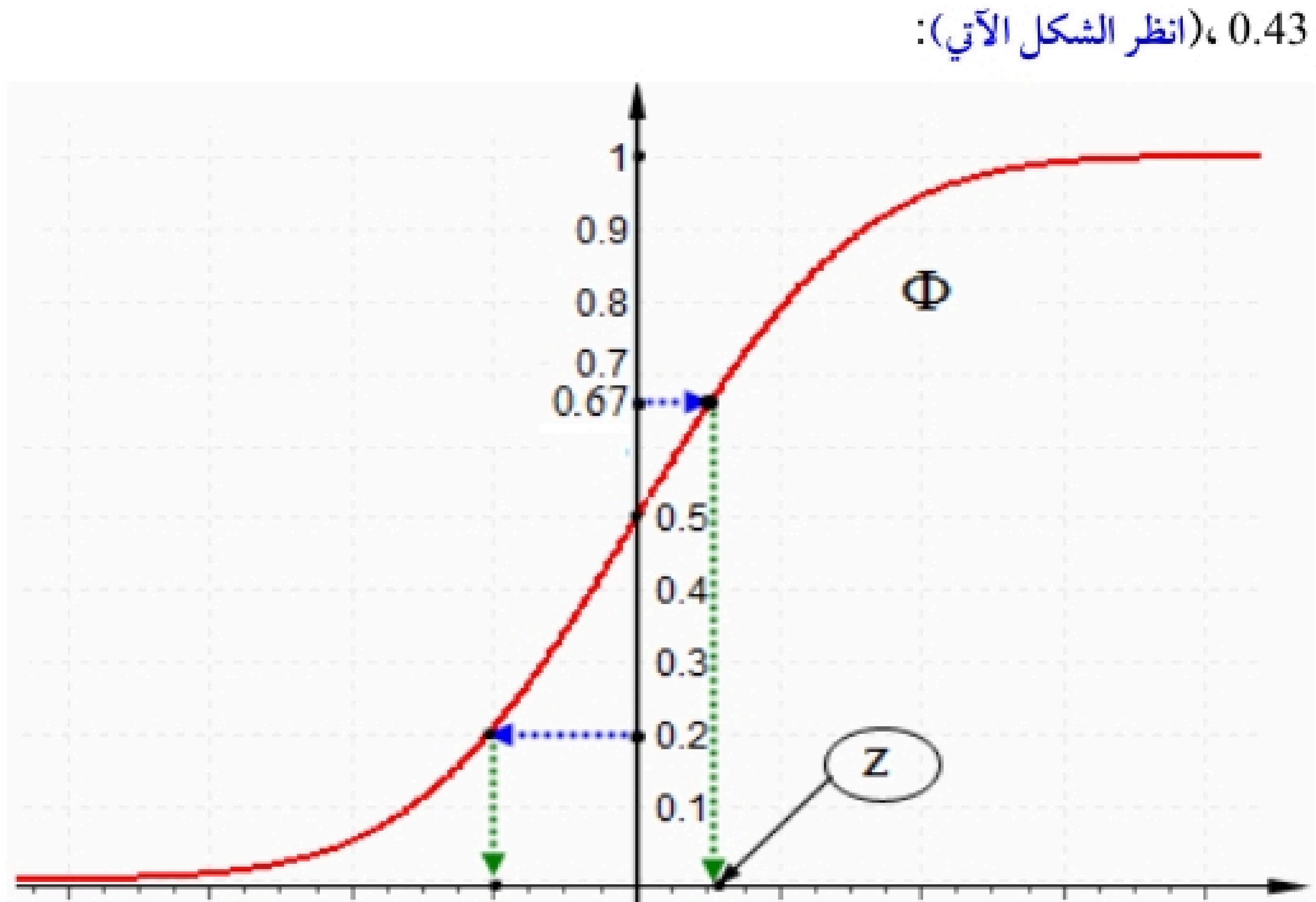
$$P(Z < z) = \Phi(z) = 0.6664$$

فعندئذ نجد من جدول التوزيع الطبيعي المعياري (القسم الخاص بالقيم الموجبة لأنَّ **0.6664** موجبة) أنَّ قيمة $z$ تساوي 0.43، وذلك من خلال جمع قيمتي z الرأسيّة والأفقية المقابلتين للقيمة 0.6664 وهما 0.03+0.4، والشكل الآتي يوضِّح لنا ذلك:

الجدول (٦,١٠)

| Z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |

نشير هنا إلى أنَّه في حال كانت قيمة الاحتمال التي لدينا غير موجودة في جدول التوزيع الطبيعي المعياري فإنَّنا نستعين بالأقرب إليها لتعيين قيمة $z$ المقابلة، وفي حال كانت قيمة هذا الاحتمال واقعة بالوسط تماماً بين قيمتين $z_1$ و $z_2$ فإنَّنا سنأخذ قيمة المتوسط لـ $z_1$ و $z_1$ كقيمة لـ $z$. كما يمكن استنتاجها بشكل تقريبي مقبول باستخدام منحنى دالة التوزيع الاحتمالية $\Phi(z)$ (في حال توفرها لدينا)، فعلى سبيل المثال نجد من أجل القيمة السابقة $P(Z < z) = 0.6664 \approx 0.67$ أنَّه لدينا من العرض البياني لدالَّة التوزيع الطبيعي المعياري، أنَّ قيمة $z$ هي 0.43، (انظر الشكل الآتي):




شكل (٦,٢٤).

وكما يوضِّح لنا الشكل السابق أنَّه إذا كان $P(Z < z) = \Phi(z) = 0.2$ فإنَّ قيمة z المقابلة تساوي تقريباً 0.8−، وبالفعل إذا رجعنا إلى جدول التوزيع الطبيعي المعياري (القسم الخاص بالقيم السالبة لـ **z**) فإنَّنا سنجد من أجل $\Phi(z) = 0.2005$ أنَّ قيمة $z$ المقابلة هي

. $z = -0.84$

نشير هنا إلى أنَّه توجد جداول عديدة ومتنوعة في طريقة عرضها والدقّة المستخدمة فيها، وقد استخدمنا هذا الجدول لبساطته وسهولة التعامل معه، ومن أجل الجداول المقدَّمة سنضع $\Phi(z) = 1$ من أجل قيم $3.50 \le z$ ، وكذلك سنضع $\Phi(z) = 0$ من أجل قيم $-3.50 \ge z$ .

الآن وبطريقة مماثلة لما سبق يمكننا تنفيذ نماذج عديدة من الحساب الاحتمالي باستخدام جدول التوزيع الطبيعي المعياري، ومنها على سبيل المثال الاحتمالات الآتية:

$$P(Z \ge 0.71) = 1 - P(Z < 0.71) = 1 - 0.7611 = 0.2389$$

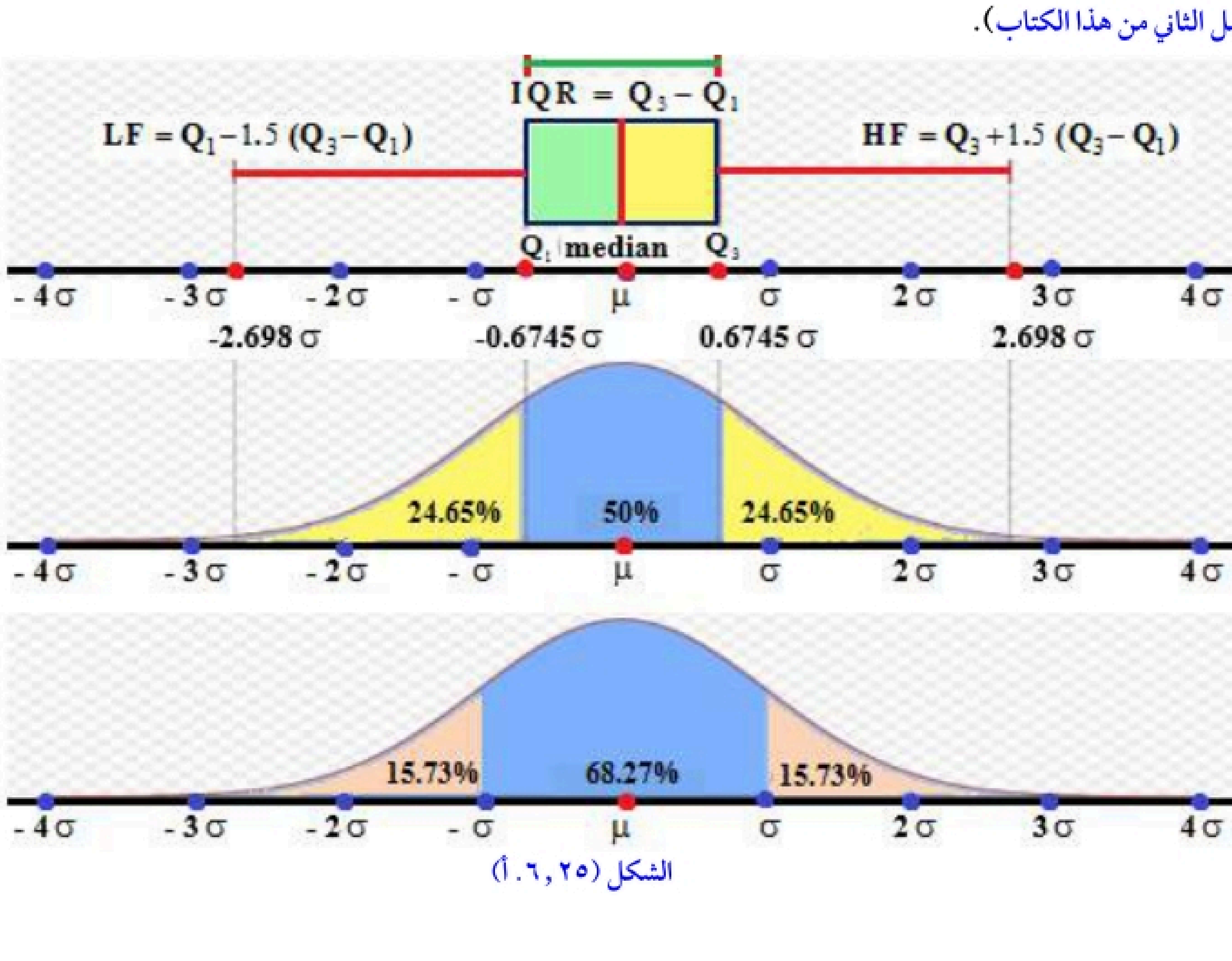
$$P(-4.10 \le Z < 1.03) = P(-4.10 < Z \le 1.03) = P(-4.10 < Z < 1.03)$$
$$= \Phi(1.03) - \Phi(-4.10) \approx 0.8485 - 0.00 = 0.8485$$

أمّا إذا كان متغيِّر عشوائي $X$ يخضع للتوزيع الطبيعي غير المعياري، فعندئذ يمكن حساب احتمالات متعلِّقة بهذا المتغيِّر العشوائي من خلال عمليّة استعياره Standardization (أي تحويله إلى متغيِّر عشوائي طبيعي معياري)، فلو كان لدينا على سبيل المثال $X \sim \mathbf{N}(\mu, \sigma)$ وطلب منّا حساب $P(X < x)$ ، فإنَّنا نقوم بعمليّة استعيار للمتغيِّر العشوائي $X$ من خلال وضع $Z := \frac{X - \mu}{\sigma}$ ، وبفرض أنَّ $z := \frac{x - \mu}{\sigma}$ فتنتج لدينا العلاقة الآتية:

$$P(X < x) = P\left(\frac{X - \mu}{\sigma} < \frac{x - \mu}{\sigma}\right) = P(Z < z) = \Phi(z)$$

وأمّا القيمة $\Phi(z)$ فإنَّها تستخرج من جدول التوزيع الطبيعي المعياري كما نوَّهنا عن ذلك سابقاً.

أخيراً نقدِّم العرض الآتي الذي يظهر لنا توزع البيانات الناتجة عن تجربة تتبع التوزيع الطبيعي $\mathbf{N}(\mu, \sigma)$ (أو لبيانات مولَّدة بوساطة التوزيع الطبيعي $\mathbf{N}(\mu, \sigma)$) وعلاقتها بالانحرافات المعيّارية والقيم المتطرفة بصغرها وكبرها (انظر الشكل الآتي واستنتج التفسيرات حسب ما ورد في الفصل الثاني من هذا الكتاب).




الشكل (٦,٢٥. أ)

علماً أنَّ الربيعيات $Q_1$ ، $Q_2$ و $Q_3$ تحسب من أجل التويع الطبيعي $N(\mu , \sigma)$ وفقاً للعلاقات الآتية:

١- قيمة الربيعي الأول $Q_1$ توافق تلك القيمة $x$ التي تكون فيها العلاقة الآتية محقَّقةً:

$$\frac{1}{4} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \, dt$$

ومن أجل من أجل التويع الطبيعي المعّياري يكون لدينا

$$0.25 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} \, dt \quad \Rightarrow \quad x = -0.6745$$

٢- قيمة الربيعي الثاني (الوسيط) $Q_2 = \tilde{x}$ توافق تلك القيمة $x$ التي تكون فيها العلاقة الآتية محقَّقةً:

$$\frac{1}{2} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \, dt$$

ومن أجل من أجل التويع الطبيعي المعّياري يكون لدينا

$$0.5 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} \, dt \quad \Rightarrow \quad x = 0$$

٣- من أجل الربيعي الثالث $Q_3$ قيمته توافق تلك القيمة $x$ التي تكون فيها العلاقة الاتية محقَّقةً:

$$\frac{3}{4} = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{(t-\mu)^2}{2\sigma^2}} \, dt$$

ومن أجل من أجل التويع الطبيعي المعّياري يكون لدينا

$$0.75 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} \, dt \quad \Rightarrow \quad x = 0.6745$$

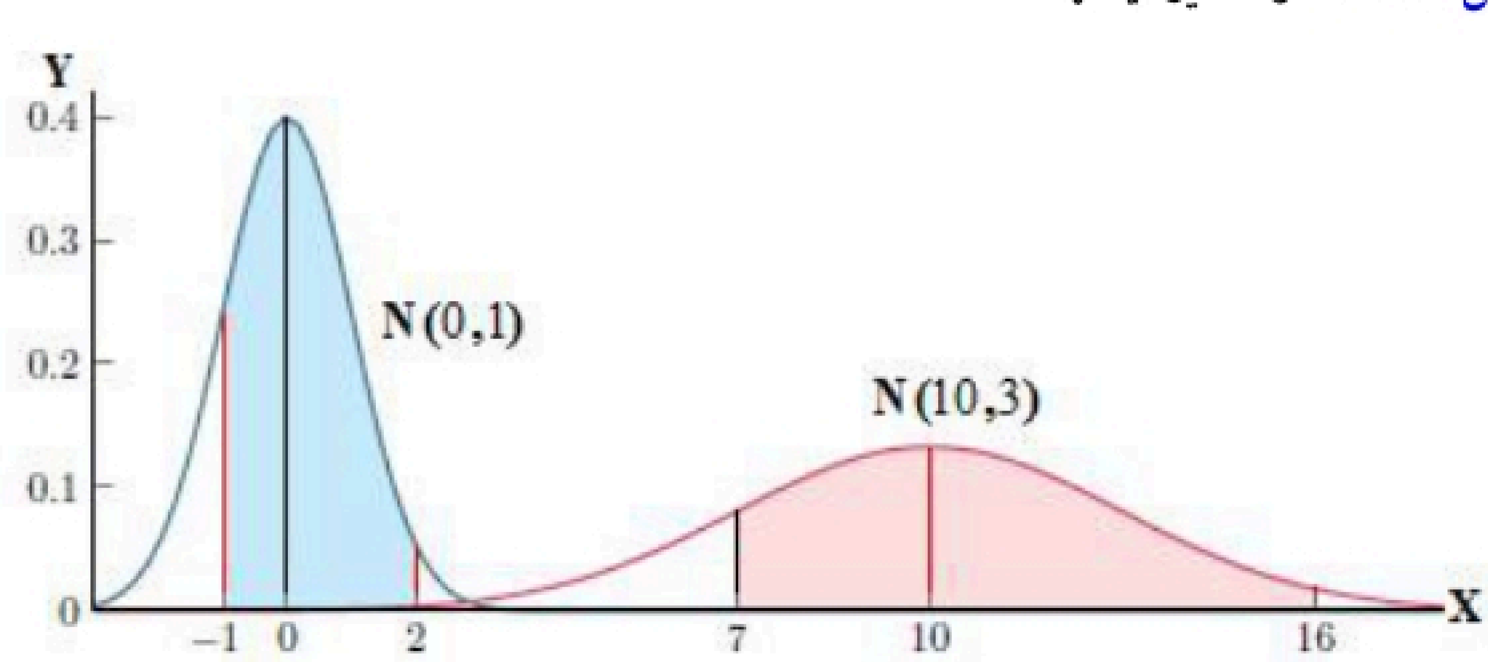
**(٦,٢,٣,٣) أمثلة**

١- لنأخذ $X \sim N(10,3)$ ، ولنقم بحساب قيمة الاحتمال $P(7 \le X < 16)$ ، ومن ثمَّ تقديم التفسير الهندسي لما يحدث نتيجة لعملية الاستعيار.

**الحـل**: من أجل ذلك لنقم أولاً باستعيار المتغيِّر العشوائي المُعطى $X$ فنجد الآتي:

$$P(7 \le X < 16) = P\left(\frac{7-10}{3} \le \frac{X-10}{3} < \frac{16-10}{3}\right) = P(-1 \le Z < 2)$$
$$= \Phi(2) - \Phi(-1) = 0.9772 - 0.1587 = 0.8185$$

والآن لو أمعنا النظر في الشكل الآتي لمعرفة ما الذي حدث للمساحة المشغولة من قبل المتغيِّر العشوائي المُعطى $X$ بعد عمليّة الاستعيار، فإنّنا نلاحظ تغيُّر شكل المساحة دون تغيُّر قيمتها.




الشكل (٦,٢٥.ب)

٢- لنأخذ $X \sim \mathbf{N}(-0.5\ ,\ 0.5)$ ، فإذا كان لهذا المتغّر العشوائي قيمة بين 0.55 - و 1.01 فما هما القيمتين اللتين تقع بينهما قيمة المتغيِّر العشوائي $Z$ الناتج عن عمليّة الاستعيار $X$ ؟

الحـل: من أجل ذلك لنقم باستعيار المتغيِّر العشوائي المُعطى $X$ فنجد الآتي:

$$P(-0.55 \le X < 1.01) = P\left(\frac{-0.55-(-0.5)}{0.5} \le \frac{X-(-0.5)}{0.5} < \frac{1.01-(-0.5)}{0.5}\right) = P(-0.10 \le Z < 3.02)$$

وهكذا نجد القيمتين اللتين تقع بينهما قيمة المتغيِّر العشوائي $Z$ الناتج عن عمليّة الاستعيار $X$ هما 0.10- و 3.02 .

٣- لنأخذ $X \sim \mathbf{N}(-1, 1.5)$ ، ولنقم بحساب قيمة الاحتمال $P(-0.55 \le X < 1.01)$ .

الحـل: من أجل ذلك لنقم أولاً باستعيار المتغيِّر العشوائي المُعطى $X$ فنجد الآتي:

$$P(-0.55 \le X < 1.01) = P\left(\frac{-0.55-(-1)}{1.5} \le \frac{X-(-1)}{1.5} < \frac{1.01-(-1)}{1.5}\right)$$

وبوضع $Z = (X+1)/1.5$ يصبح لدينا:

$$P(-0.55 \le X < 1.01) = P(0.30 \le Z < 1.34) = \Phi(1.34) - \Phi(0.30) = 0.9099 - 0.6179 = 0.2920$$

٤- لنأخذ $X \sim \mathbf{N}(1,2)$ ، ولنقم بحساب الاحتمالات الآتية:

**a)** $P(X < 1.53)$  **b)** $P(X > -2.74)$  **c)** $P(-2.74 < X \le 1.53)$

الحـل: بما أنَّ المتغيِّر العشوائي المُعطى $X$ ليس معيارياً فإنَّنا سنقوم باستعياره في كل حالة من الحالات الثلاث المُطلوب حساب احتمالها، فنجد من أجل الطلب:

**a)** أنّه باستخدام جدول التوزيع الطبيعي المعّياري لتعيين قيمة الاحتمال المطلوب ما يلي:

$$P(X < 1.53) = P\left(\frac{X-1}{2} < \frac{1.53-1}{2}\right) = P(Z < 0.2650) = \Phi(0.2650) = 0.6045$$

**b)** أنّه وبشكل مماثل لما قمنا به في الطلب السابق، سيكون لدينا:

$$\begin{aligned} P(X > -2.74) &= 1 - P(X \le -2.74) = 1 - P(X < -2.74) = 1 - P\left(\frac{X-1}{2} < \frac{-2.74-1}{2}\right) \\ &= 1 - P(Z < -1.8700) = 1 - \Phi(-1.8700) = 1 - 0.0307 = 0.9693 \end{aligned}$$

**c)** أنّه وبشكل مماثل لما قمنا به في الطلبين السابقين، سيكون لدينا:

$$\begin{aligned} P(-2.74 \le X < 1.53) &= P\left(\frac{-2.74-1}{2} < \frac{X-1}{2} \le \frac{1.53-1}{2}\right) = P\left(\frac{-2.74-1}{2} \le \frac{X-1}{2} < \frac{1.53-1}{2}\right) \\ &= 1 - P(-1.8700 \le Z < 0.2650) = \Phi(0.2650) - \Phi(-1.8700) \\ &= 0.6045 - 0.0307 = 0.5738 \end{aligned}$$

٥- لنأخذ $X \sim \mathbf{N}(-2, 1.5)$ ، ولنقم بتعيين قيمة $x$ في كل من الحالات الآتية:

**a)** $P(X < x) = 0.9750$  **b)** $P(X > x) = 0.8549$

الحـل: بما أنَّ المتغيِّر العشوائي المُعطى $X$ ليس معيارياً فإنَّنا سنقوم باستعياره في كل حالة من الحالتين المُطلوب حساب احتمالها،

فنجد من أجل الطلب:

**a**) أنَّه باستخدام جدول التوزيع الطبيعي المعّياري لتعيين قيمة الاحتمال المطلوب ما يلي:

$$P(X < x) = P\left(\frac{X - (-2)}{1.5} < \frac{x - (-2)}{1.5}\right) = P\ (Z < z\ ) = \Phi(z) = 0.9750$$

ومن جدول التوزيع الطبيعي المعياري نجد أنَّ قيمة $z$ الموافقة للاحتمال 0.9750 هي 1.96، ومن ثمَّ نحسب قيمة $x$ المطلوبة على النحو الآتي:

$$z = \frac{x + 2}{1.5} = 1.96 \ \Rightarrow\ x = 2.94 - 2 = 0.94$$

**b**) أنَّه وبشكل مماثل لما قمنا به في الطلب السابق سيكون لدينا:

$$\begin{aligned} P\ (X > x\ ) &= 1 - P\ (X \le x\ ) = 1 - P\ (X < x\ ) \\ &= 1 - P\left(\frac{X - (-2)}{1.5} < \frac{x - (-2)}{1.5}\right) \\ &= 1 - P\ (Z < z\ ) = 1 - \Phi(z\ ) = 0.0505 \end{aligned}$$

ومن هذه العلاقة الأخيرة نجد أنَّ $P\ (Z < z\ ) = 1 - 0.0505 = 0.9495$، ومن جدول التوزيع الطبيعي المعّياري نجد القيمة $z$ الموافقة للاحتمال 0.9495 هي 1.64، ومنه يكون لدينا $z = \frac{x + 2}{1.5} = 1.64$، ومن ثمَّ نجد أنَّ $x = 2.46 - 2 = 0.46$ .

من الأسماء المرتبطة بالتوزيع الطبيعي ما يُعرف باسم **التوزيع الطبيعي اللوغارتمي** Lognormal Distribution، وهو توزيع يعتمد في تعريفه على التوزيع الطبيعي، حيث يُقال عن متغيِّر عشوائي $Y$ إنَّه خاضع للتوزيع الطبيعي اللوغارتمي إذا كانت دالَّة كثافته الاحتمالية $f_Y$ مُعطاة من خلال العلاقة الآتية:

$$f_Y\ (y\ ) = \frac{1}{\sqrt{2\pi}\ \sigma y} \mathbf{exp}\left[-\frac{(\mathbf{ln}\, y - \mu)^2}{2\ \sigma^2}\right] \cdot \mathbf{I}_{(0\ ,\ \infty)}(y\ )$$

فنلاحظ أنَّ لهذا التوزيع مَعْلمتين هما $\mu \le 0$ و $\sigma > 0$ يُدعيان المتوسط والانحراف المعياري لهذا التوزيع أيضاً (**ولذلك سنرمز لهذا التوزيع بـ** ($\mathbf{LN}(\mu,\sigma)$). أمّا الرسم البياني لدالَّة كثافته الاحتمالية $f_Y$ فلها العرض في الشكل الآتي (٢٦ ,٦. أ)، حيث نلاحظ أنَّ هنا أنَّ قيمة الانحراف المعياري هي المسؤولة عن انسحاب ذروة دالَّة الكثافة الاحتمالية يمناً ويُساراً، فكلما كبرت قيمة $\sigma$ فإنَّ ذلك سيؤدي إلى ازياح أكبر لذروة دالَّة الكثافته الاحتمالية نحو اليسار. كما يلاحظ أنَّه بفرض لـ $X$ توزيع طبيعي $\mathbf{N}(\mu,\sigma)$ فإنَّ المتغيّر العشوائي $Y$ الموصوف أعلاه له العرض $Y = \mathbf{ln}\, X$.

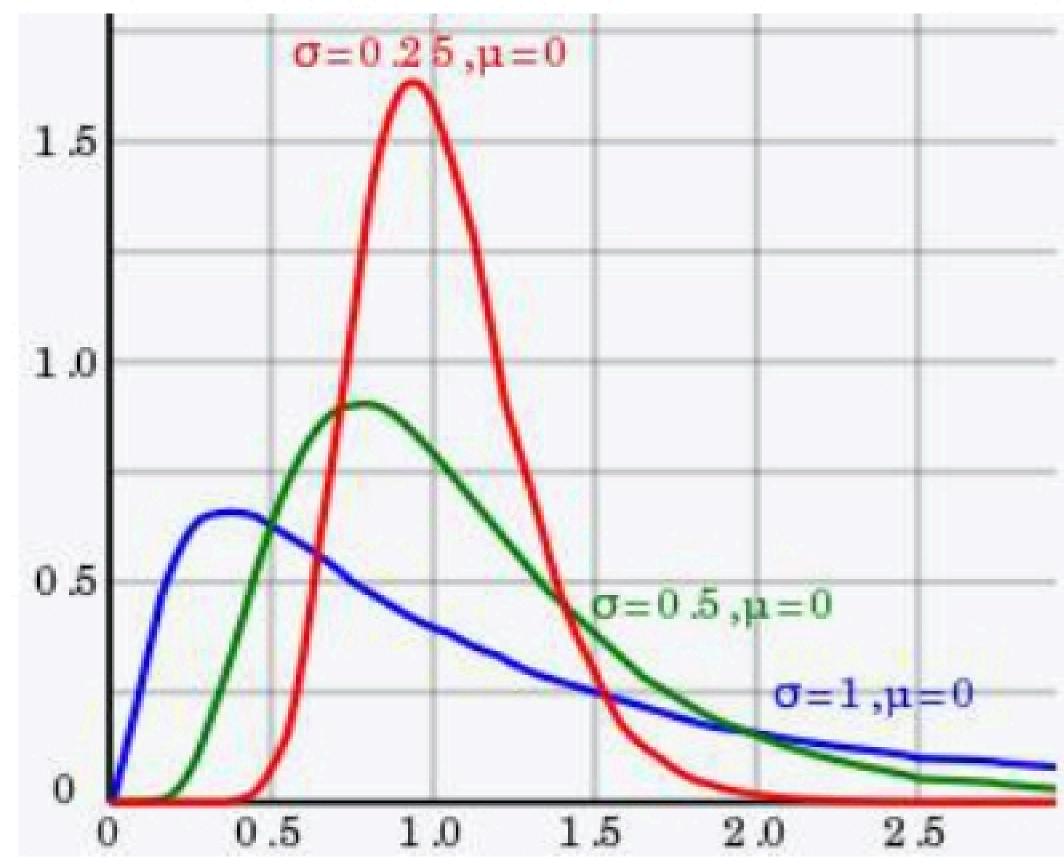
في الحالة الخاصة عندما تكون $\mu = 0$ و $\sigma = 1$ فإنَّه يصبح للعلاقة السابقة العرض الآتي:

$$f_Y\ (x\ ) = \frac{1}{\sqrt{2\pi}\ y} \mathbf{exp}\left[-\frac{(\mathbf{ln}\, x\ )^2}{2}\right] \cdot \mathbf{I}_{(0\ ,\ \infty)}(x\ )$$

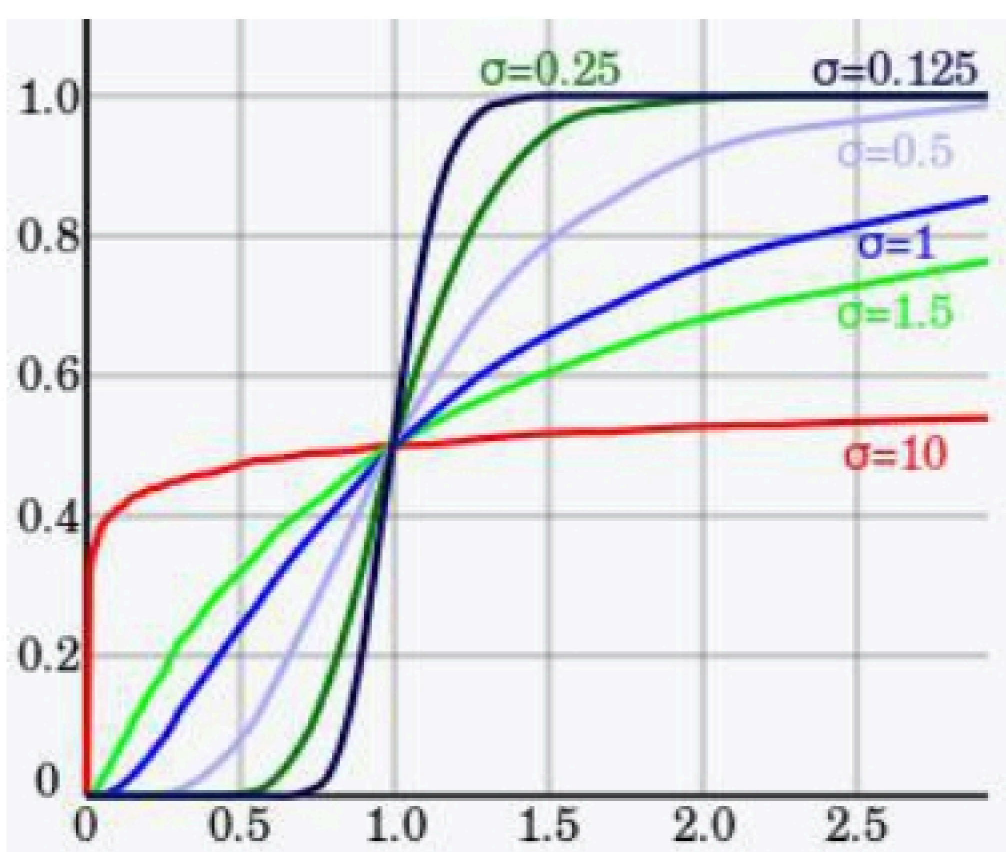
و أمّا دالَّة التوزيع الاحتمالية لـ $Y$ فلها العرض الآتي من أجل كل $x > 0$ :

$$F_Y\ (x\ ) = \Phi\left(\frac{\mathbf{ln}(x\ ) - \mu}{\sigma}\right) = \frac{1}{\sigma_X \sqrt{2\pi}} \int_{-\infty}^{\mathbf{ln}\, x} \mathbf{e}^{-\frac{(t - \mu_X\ )^2}{2\ \sigma_X^2}}\ dt$$

ولرسمها البياني العرض الآتي في الشكل (٦,٢٦. ب).




الشكل (٦,٢٦. أ)




الشكل (٦,٢٦. ب)

لهذا التوزيع تطبيقات كثيرة في مجالات علمية متعدّدة منها العلوم الهندسية والفيزيائية والكيميائية و الاجتماعية، حيث يلعب التوزيع اللوغاريتمي دورًا مهمًا في التصميم الاحتمالي لأنَّ القيم السالبة لبعض الظواهر تكون في بعض الأحيان مستحيلة ماديًا. من استخداماته نمذجة فشل الاجهاد، ومعدلات الفشل ، وظواهر أخرى تنطوي على مجموعة كبيرة من البيانات.

### (٦,٢,٤) توزيع ستودنت Student Distribution

فيما يلي سنتعرف على توزيع له أهمية خاصّة في مجال الإحصاء الرياضياتي، وعلى وجه الخصوص في الدراسات التي تتعلّق بتقديرات واختبارات ذات صلة بمتوسط مجتمع إحصائي. لقد اكتشف الكيميائي والرياضياتي الإنكليزي **غوسيت** William Sealy Gosset (1876 –1937) هذا التوزيع عام 1908 في أثناء عمله في مصنع غينيس للجُعّة، وفي ذلك الحين لم يكن يسمح مصنع غينيس لموظفيه بنشر أعمالهم، ولذلك استخدم غوسيت اسماً مستعارًا هو "**طالب** Student"، وبناءً على ذلك عرف هذا التوزيع باسم توزيع **ستودنت** (كما أنَّه يُعرف باسم توزيع–t **t-Distribution** أيضاً).

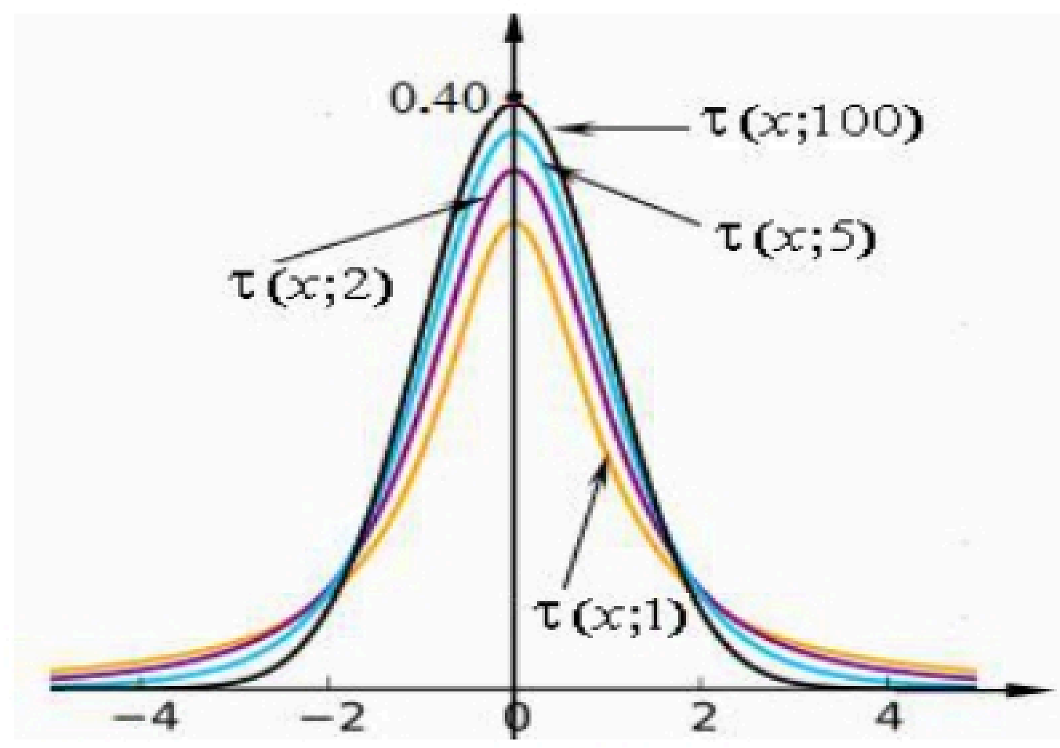
#### (٦,٢,٤,١) تعريفه

ليكن $X$ متغيِّراً عشوائياً مستمراً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ ، فعندئذ يُقال عن $X$ إنَّه خاضع لتوزيع **ستودنت** بمَعْلمة $n \in \mathbb{N}$ تُدعى **عدد درجات الحرية** Degrees of Freedom إذا كانت دالّة كثافته الاحتمالية $f_X$ مُعطاة من خلال العلاقة الآتية:

$$f_X(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)\cdot\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n\pi}\cdot\Gamma\left(\frac{n}{2}\right)} \quad ;\forall x \in \mathbb{R} \qquad [6,13]$$

علماً أنَّ $\Gamma(x)$ ترمز لدالّة غاما عند القيمة $x$ (**انظر الملحق A**).

إنَّ الرسم البياني لدالّة الكثافة الاحتمالية لتوزيع ستودنت نجده من أجل قيم مختلفة لعدد درجات الحرية $n$ كما في الشكل الآتي:




الشكل (٦.٢٧ ,أ)

**(٦,٢,٤,٢) ملاحظات**

١- سنرمز لتوزيع هـذا المتغيِّر العشوائي بـ $T(n)$ ، وأمّا دالَّة كثافته الاحتمالية فسوف نرمز لها بـ $\tau(x;n)$ .

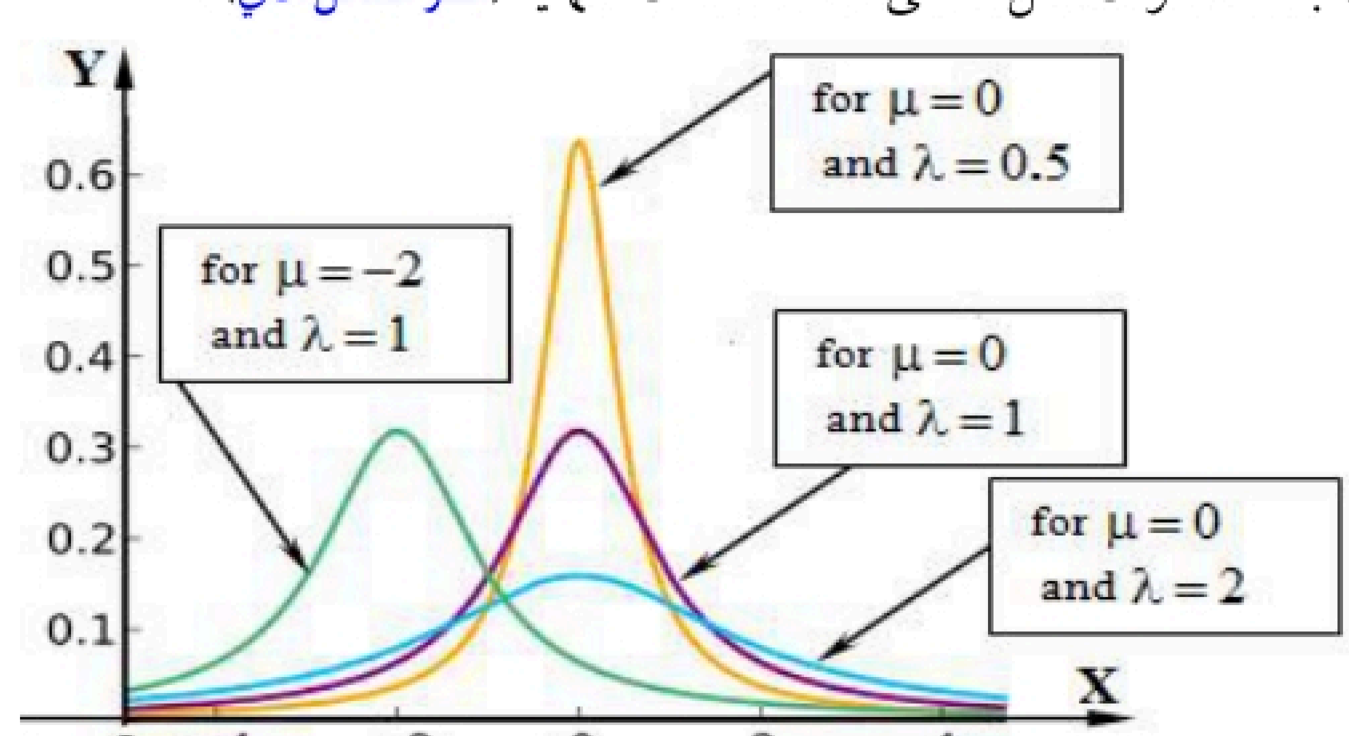
٢- في الحالة الخاصَّة عندما $n=1$ يصبح لدالَّة الكثافته الاحتمالية $f_X$ العرض الآتي:

$$f_X(x) = \frac{1}{\pi(1+x^2)} \quad ;\forall x \in \mathbb{R} \qquad [6,14\text{-}a]$$

وهذه الدالَّة الأخيرة هي دالَّة الكثافته الاحتمالية لما يُعرف باسم **توزيع كوشي** Cauchy Distribution ذو المَعْلمتين $\mu = 0$ و $\lambda = 1$ (نسبة إلى الرياضياتي الفرنسي كوشي **Augustin-Louis Cauchy (1789-1857)**). علماً أنَّ الصيغة العامّة لدالَّة الكثافته الاحتمالية لتوزيع كوشي ذي المَعْلمتين $\mu \in \mathbb{R}$ و $\lambda > 0$ هي:

$$f_X(x) = \frac{\lambda}{\pi\left(\lambda^2 + (x-\mu)^2\right)} \quad ;\forall x \in \mathbb{R} \qquad [6,14\text{-}b]$$

وقيم $\mu$ هنا تلعب دور مُعَامل الانسحاب لمنحنى دالَّة الكثافة الاحتمالية يُمنة أويُسرة فوق المحور **oX** في حين أنَّ قيم $\lambda$ تلعب دور مُعَامل التدبب للأعلى أوالانبساط نحو الأسفل لمنحنى دالَّة الكثافة الاحتمالية (انظر الشكل الآتي).




الشكل (٦.٢٧ ,ب)

وسنرمز لهذا التوزيع بـ $C(\mu,\lambda)$ ، وأما دالَّة كثافته الاحتمالية فسوف نرمز لها بـ $c(x;\mu,\lambda)$ .

٣- إنَّ توزيع ستودنت يُستخدم في اختبارات عديدة منها اختبارات القيم الوسطى، وفي تحليل الانحدار.

٤- باستخدام تعريف دالَّة التوزيع الاحتمالية لمتغيِّر عشوائي مستمرّ يكون لدالَّة توزيع هذا المتغيِّر العشوائي العرض الآتي:

$$F_X(x)=\int_{-\infty}^{x}\frac{\Gamma\left(\frac{n+1}{2}\right)\cdot\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n\pi}\;\Gamma\left(\frac{n}{2}\right)}\,dt \qquad ;\forall x\in\mathbb{R}$$

والتي يمكن صياغتها باستخدام الدالَّة فوق الهندسية $H(x)$ على النحو الآتي أيضاً:

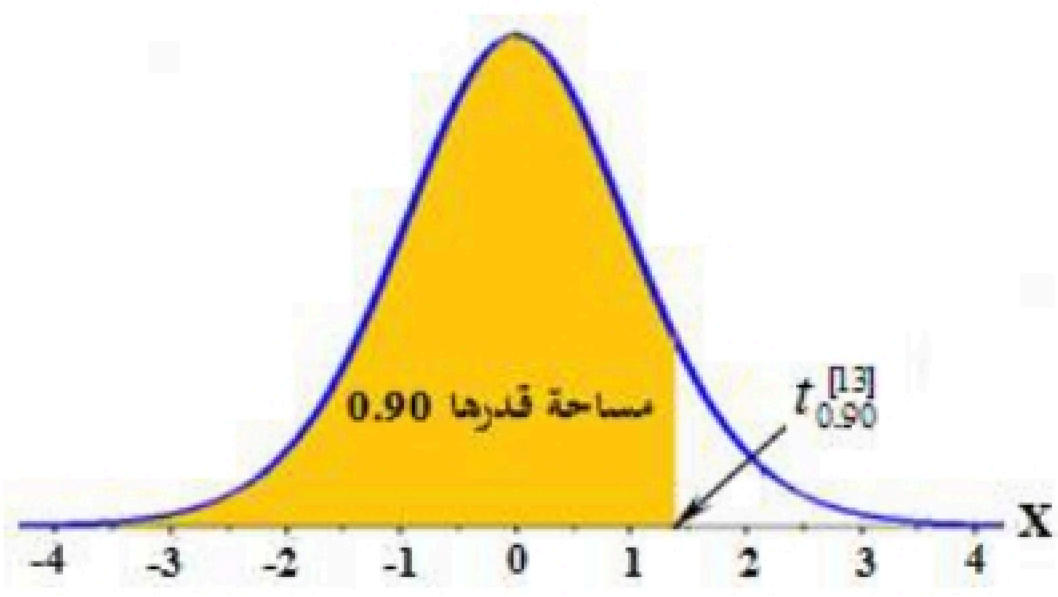
$$F_X(x)=\frac{1}{2}+x\cdot\Gamma\left(\frac{n+1}{2}\right)\frac{2H\left(\frac{1}{2},\frac{n+1}{2};\frac{3}{2};-\frac{x^2}{n}\right)}{\sqrt{n\pi}\;\Gamma\left(\frac{n}{2}\right)}$$

نشير هنا إلى أنَّه من الصعب جداً (كما هو واضح) التعامل مع هذه الصيغة في الحسابات، ولذلك تُستخدم جداول لحساب قيم دالَّة التوزيع السابقة، وقد قُدِّم أحد هذه الجداول في آخر هذا الكتاب، وهذا الجدول يعطينا قيماً لهذه الدالَّة من أجل قيم مختلفة لدرجات الحرية $n$ وقيمًا مُحدَّدة للاحتمالات التي تأخذ عندها دالَّة التوزيع هذه القيم، ومنها **0.999 0.995 0.99 0.975 0.95 0.90**، وكذلك توجد نماذج أخرى من الجداول تقدّم قيمًا للاحتمالات من قبيل **0.001 0.005 0.01 0.025 0.05 0.10**، فعلى سبيل المثال من أجل $n=5$ و $P(X<x)=0.975$ فإنَّ قيمة $x$ الموافقة لهذا الاحتمال تساوي 2.571، بمعنى أنَّ $x$ هنا هي تلك القيمة على المحور **oX** التي يقع على يسارها وتحت منحنى دالَّة الكثافة الاحتمالية لتوزيع ستودنت ذي الـ 5 درجات حرية، مساحة قدرها 0.975 (الشكل الآتي يوضِّح لنا ذلك). كذلك نجد من أجل $n=8$ أنَّ $P(X<3.355)$ تساوي 0.995.

الجدول (٦,١١)

| α | 0.90 | 0.95 | 0.975 | 0.99 | 0.995 | 0.999 |
|---|---|---|---|---|---|---|
| n | | | | | | |
| 1. | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 | 318.313 |
| 2. | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 22.327 |
| 3. | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 10.215 |
| 4. | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 7.173 |
| 5. | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 5.893 |
| 6. | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 5.208 |
| 7. | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.782 |
| 8. | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 4.499 |
| 9. | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 4.296 |

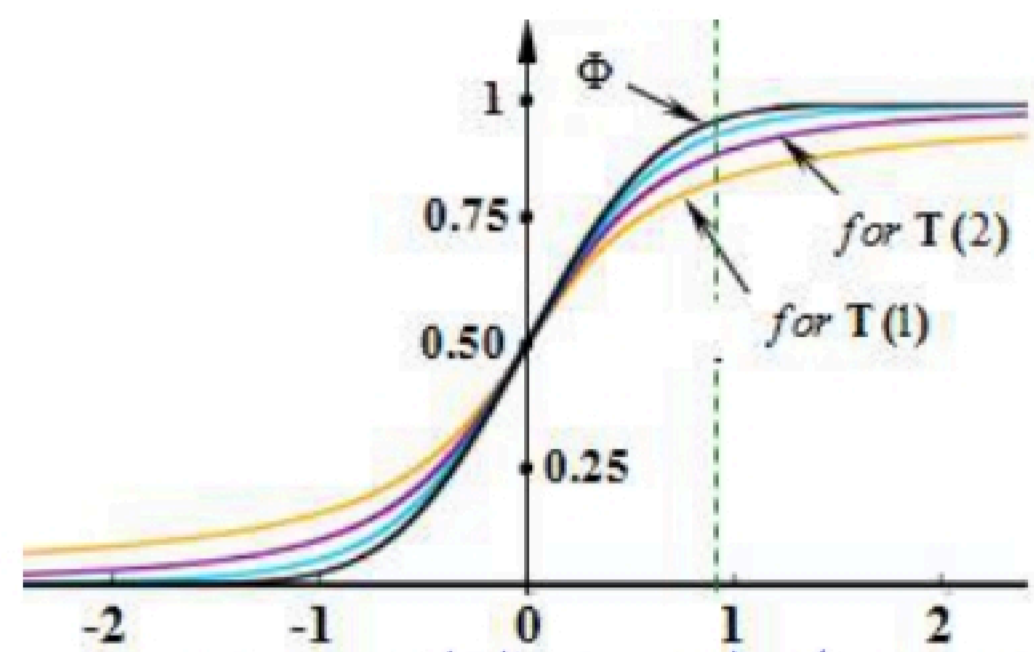
٥- سوف نرمز بـ $t_{\alpha}^{[n]}$ للقيمة $t$ الواقعة على المحور الأفقي **oX** التي يقع على يسارها وتحت منحنى الكثافة الاحتمالية $\tau(x;n)$ مساحة قدرها α (قيمة هذه المساحة تساوي $P(X<x)$)، فعلى سبيل المثال من أجل $n=13$ و $\alpha=0.90$ يكون لدينا $t_{0.90}^{[13]}=1.350$، أي إنَّ $P(X<1.350)=0.90$ .



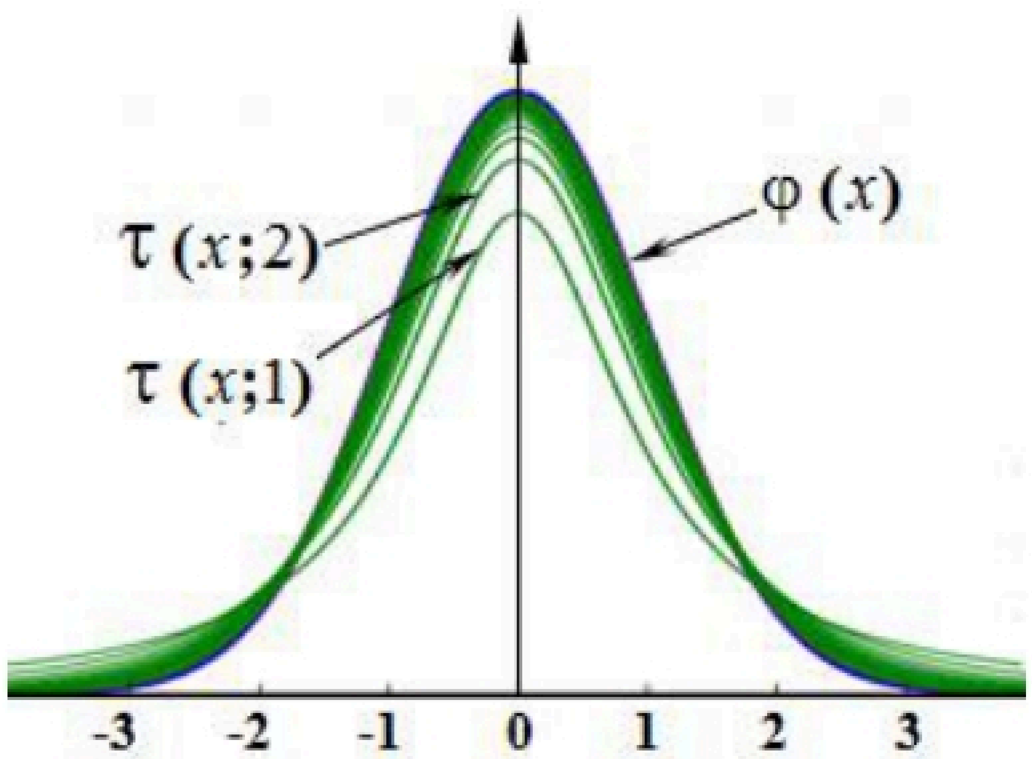

الشكل (٦,٢٧.ج)

٦- لو أمعنَّا النظر في جدول التوزيع الطبيعي المعياري وجدول توزيع ستودنت لوجدنا أنَّه من أجل قيم درجات الحرية $n\geq 32$

يصبح الفرق بين قيم دالتي التوزيع الاحتمالية لهذين التوزيعين ضئيلاً (انظر الشكل الآتي) حيث يُلاحظ تقارب شكل دالَّة توزيع ستودنت من دالَّة التوزيع الطبيعي المعياري مع تزايد قيم عدد درجات الحرية $n$ . كما يعرض الشكلين الآتيين الوضع النسبي لمنحنيات دوال الكثافة الاحتمالية، وكذلك دوال التوزيع الاحتمالية، لتوزيع ستودنت لدى تزايد قيم عدد درجات الحرية $n$ ، ومقارنتها مع التوزيع الطبيعي.




الشكل (٦,٢٨ .أ) دوال توزيع احتمالية لتوزيع ستودنت




الشكل (٦,٢٨ .ب) دوال كثافة احتمالية لتوزيع ستودنت

فعلى سبيل المثال لو أخذنا $X \sim \mathbf{N}(0,1)$ فإنَّنا سنجد $P(X < 0.96) = 0.8315$ ، ولقد اخترنا القيمة $x = 0.96$ لأنَّ الفروق بين قيم دوال التوزيع عند هذه النقطة كبيرة نسبياً (انظر الشكل السابق)، فلو قمنا باستخدام الحاسوب لحساب المقادير الآتية المتعلّقة بقيم التوزيعات $\mathbf{T}(n)$ (لأنَّ بعض قيمها عند النقطة المذكورة آنفًا قد تقع خارج محتويات الجدول المخصص لهذا التوزيع) فإنَّنا سنجد ما يلي:

الجدول (٦,١٢)

| من أجل | لدينا | | من أجل | لدينا |
|---|---|---|---|---|
| $X \sim \mathbf{T}(1)$ | $P(X < 0.96) = 0.7435$ | | $X \sim \mathbf{T}(50)$ | $P(X < 0.96) = 0.8292$ |
| $X \sim \mathbf{T}(2)$ | $P(X < 0.96) = 0.7808$ | | $X \sim \mathbf{T}(75)$ | $P(X < 0.96) = 0.8299$ |
| $X \sim \mathbf{T}(5)$ | $P(X < 0.96) = 0.8094$ | | $X \sim \mathbf{T}(100)$ | $P(X < 0.96) = 0.8303$ |
| $X \sim \mathbf{T}(25)$ | $P(X < 0.96) = 0.8269$ | | $X \sim \mathbf{T}(1000)$ | $P(X < 0.96) = 0.8314$ |

وهكذا نلاحظ تضاؤل قيمة الفرق بينهما كلما كبرت قيمة $n$، ولهذا السبب يقوم البعض باستخدام جدول التوزيع الطبيعي لحساب احتمالات تتعلّق بتوزيع ستودنت عندما يكون عدد درجات الحرية $n \geq 32$ ، وهناك آخرون يستخدمون هذه العمليّة من أجل قيم $n \geq 27$ ، ولكن يفضَّل دوماً استخدام جدول توزيع ستودينت في حال توفره عند القيم المطلوب حساب احتمالاتها.

بخصوص الأمثلة والتطبيقات على هذا التوزيع سوف نقدِّمها لاحقاً في البحوث الإحصائية لدى تقديم فترات الثقة واختبار الفرضيات الإحصائية.

### (٦,٢,٥) توزيع كاي مربَّع Chi Square Distribution

إنَّ توزيع كاي مربَّع يُعدّ من التوزيعات المهمَّة التي لها دور مميَّز في مجال الإحصاء الرياضياتي أيضاً، وعلى وجه الخصوص في الدراسات التي تتعلَّق بتقدير واختبار فرضية إحصائية ذات صلة بتباين مجتمع طبيعي، وكذلك في مجال اختبار جودة التوفيق لتوزيع، وبعض الاختبارات غير المعلميّة الأخرى التي سنقدِّم بعضاً منها لاحقاً في الفصول التابعة للإحصاء الرياضياتي.

**(٦،٢،٥،١) تعريفه**

ليكن $X$ متغيِّراً عشوائياً مستمراً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ ، فعندئذٍ يُقال عن $X$ إنَّه خاضع لتوزيع **كاي مربَّع** بمَعْلمة $n \in \mathbb{N}$ تُدعى **عدد درجات الحرية** للتوزيع، إذا كانت دالَّة كثافته الاحتمالية $f_X$ معطاة من خلال العلاقة الآتية:

$$f_X(x) := \left(\Gamma\left(\frac{n}{2}\right)\right)^{-1}\left(\frac{1}{2}\right)^{\frac{n}{2}} x^{\frac{n}{2}-1} \mathbf{e}^{-\frac{x}{2}} \cdot \mathbf{I}_{(0,\infty)}(x) \qquad ;\forall x \in \mathbb{R} \qquad [6,15]$$

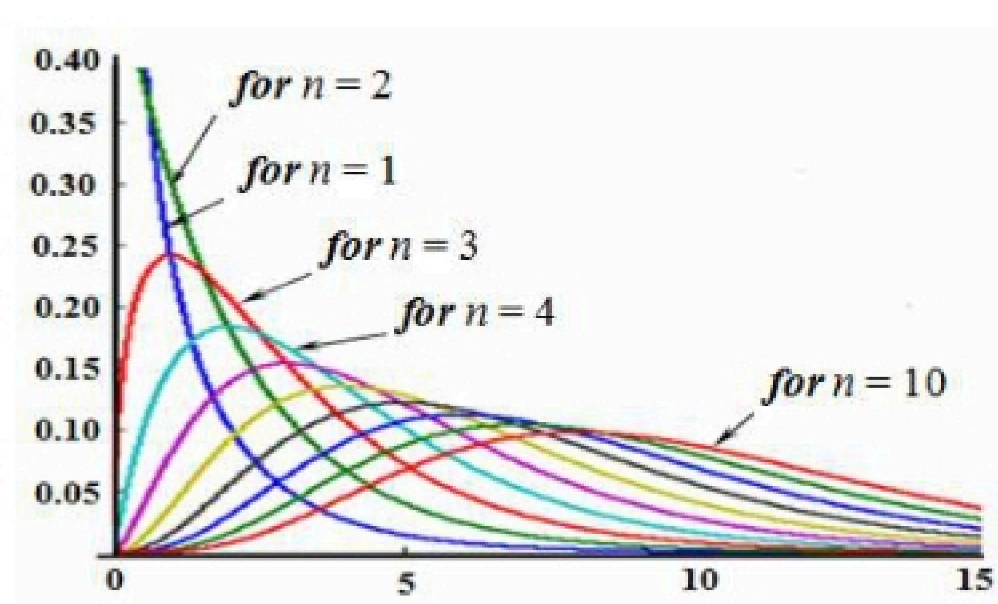
والشكل الآتي (٦،٢٨ . أ) يعرض لنا رسوماً بيانية لهذه الدالَّة من أجل قيم مختلفة لعدد درجات الحرية $n$ .

**(٦،٢،٥،٢) ملاحظات**

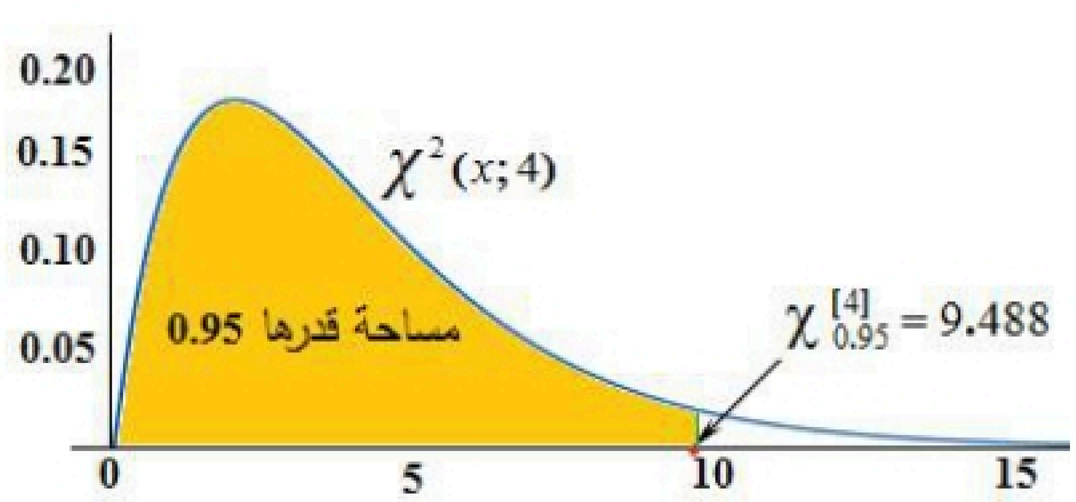
١- سنرمز بـ $\chi^2(n)$ لتوزيع هذا المتغيِّر العشوائي، وبـ $\chi^2(x;n)$ لدالَّة كثافته الاحتمالية.

٢- سنرمز بـ $\chi_{\alpha}^{[n]}$ للقيمة على المحور الأفقي **oX** والتي يقع على يسارها وتحت منحنى دالَّة الكثافة الاحتمالية لهذا المتغيِّر العشوائي مساحة قدرها $\alpha$ (**انظر الشكل (٦،٢٩ ب) حيث يوضِّح لنا ذلك**).

٣- يستخدم هذا النوع من المتغيِّرات العشوائية في اختبارات إحصائية عديدة منها اختبار فرضية إحصائية ذات صلة بتباين مجتمع طبيعي، وكذلك في مجال اختبارات لفرضيات غير مَعْلميّة مثل اختبار جودة التوفيق لتوزيع ما، واختبارات التجانس، واختبارات الاستقلال بين الظواهر.




شكل (٦،٢٩ . أ) دالَّة الكثافته الاحتمالية من أجل قيم مختلفة لـ $n$ .

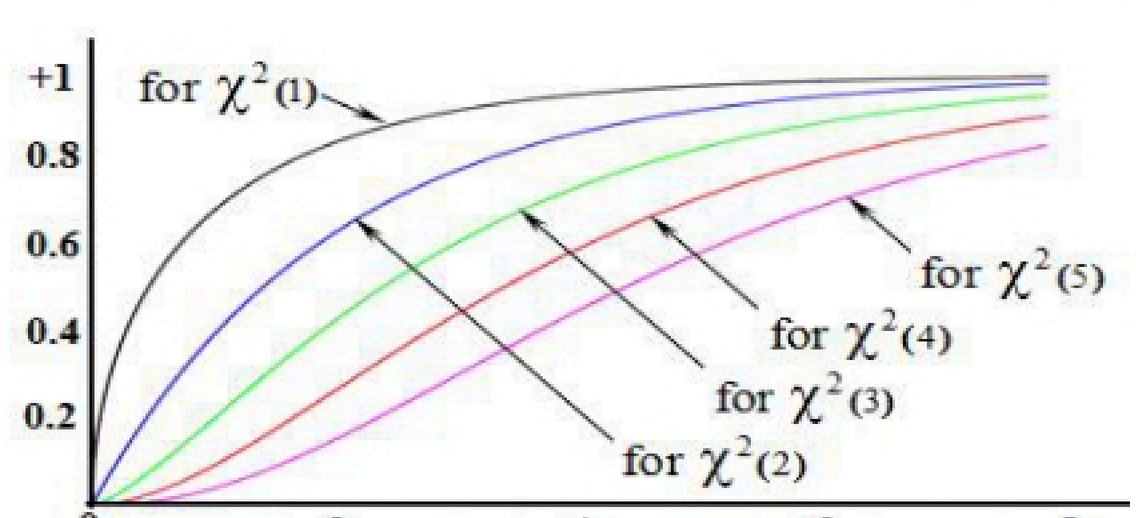



شكل (٦،٢٩ . ب)

٤- باستخدام تعريف دالَّة التوزيع الاحتمالية لمتغيِّر عشوائي مستمرّ يكون لدالَّة توزيع هذا المتغيِّر العشوائي العرض الآتي.

$$F_X(x) = \int_{-\infty}^{x} \frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)} t^{\frac{n}{2}-1} \mathbf{e}^{-\frac{t}{2}} \mathbf{I}_{(0,+\infty)}(t)\, dt \qquad ;\forall x \in \mathbb{R}$$

والشكل الآتي يقدِّم رسمها البياني من أجل قيم مختلفة لعدد درجات الحرية $n$ .




الشكل (٦،٣٠)

نشير هنا أيضاً إلى أنَّه من الصعب جداً التعامل مع هذه الصيغة في الحسابات، وكذلك يصعب عرضها من خلال دوال بسيطة، فعلى سبيل المثال يمكن تقديمها من خلال العلاقة الاتية:

$$F_X(x) = \frac{\gamma\left(\frac{n}{2}, \frac{x}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} = G\left(\frac{n}{2}, \frac{x}{2}\right)$$

علماً أنَّ $\gamma(.,.)$ في الصيغة السابقة هي دالَّة غاما غير المكتملة الدنيا وأمَّا $G(.,.)$ فإنَّها ترمز إلى دالَّة غاما المُعمَّمة (وهي نوعان خاصّان من الدوال). لذلك تُستخدم جداول خاصّة لحساب قيم دالَّة التوزيع السابقة، وقد قُدِّم أحد هذه الجداول في آخر هذا الكتاب، وهذا الجدول يُعطينا قيماً لهذه الدالَّة من أجل قيم مختلفة لدرجات الحرية $n$ وقيم مُحدَّدة للاحتمالات التي تأخذ عندها دالَّة التوزيع هذه القيم، ومنها:

**0.005 0.025 0.05 0.90 0.95 0.975 0.99 0.995**

فعلى سبيل المثال:

أ- من أجل $n = 5$ و $P(X < x) = 0.975$ فإنَّ قيمة $x$ الموافقة لهذا الاحتمال تساوي 2.571.

ب- من أجل $n = 8$ فإنَّ $P(X < 3.355)$ تساوي 0.995.

ج- من أجل $n = 4$ لدينا $P(X < 48.602) = 0.95$.

علماً أنَّ عملية استخراج القيم $x$ من جدول توزيع كاي مربَّع تتمّ بطريقة مماثلة لتلك التي استخدمت في استخراج القيم من جدول توزيع ستودنت السابق شرحها.

٥-إذا كانت $n \geq 32$ فعندئذ يمكن استخدام جدول التوزيع الطبيعي عوضاً عن جدول توزيع كاي مربع من خلال حساب المقدار $\chi_\alpha^{[n]}$ وفقاً للعلاقة الآتية:

$$\chi_\alpha^{[n]} = \frac{\left(z_\alpha + \sqrt{2n-1}\right)^2}{2} \qquad [6,15]$$

فعلى سبيل المثال لو أخذنا $\alpha = 0.975$ فإنَّنا سنجد من جدول التوزيع الطبيعي المعيّاري أنَّ $z_{0.975} = 1.96$، وكذلك لدينا من أجل $n = 50$ و $n = 100$ على الترتيب ما يلي:

$$\frac{\left(z_{0.975} + \sqrt{100-1}\right)^2}{2} = 71.903 \quad \text{and} \quad \frac{\left(z_{0.975} + \sqrt{200-1}\right)^2}{2} = 129.07$$

ومن جدول توزيع كاي مربَّع نجد $\chi_{0.975}^{[50]} = 71.420$ و $\chi_{0.975}^{[100]} = 129.65$ حيث يُلاحظ أنَّ الفارق في كلا الحالتين قريب من النصف.

### (٦, ٢, ٦) توزيع فيشر Fisher distribution

إنَّ توزيع فيشر يُعدّ من التوزيعات الاحتمالية المهمَّة جداً في الدراسات الإحصائية أيضاً، وقد كان الإحصائي الإنكليزي فيشر Sir Ronald Aylmer Fisher (1890 –1962) هو أول من وصَّف هذا التوزيع عام 1924 في عمل علميّ قام بنشره آنذاك.

#### (١, ٦, ٢, ٦) تعريفه

ليكن $X$ متغيِّراً عشوائياً مستمراً فوق فضاء احتمالي $[\Omega, \mathcal{A}, P]$ ، فعندئذ يُقال عن $X$ إنَّه خاضع لتوزيع فيشر (أو توزيع F)

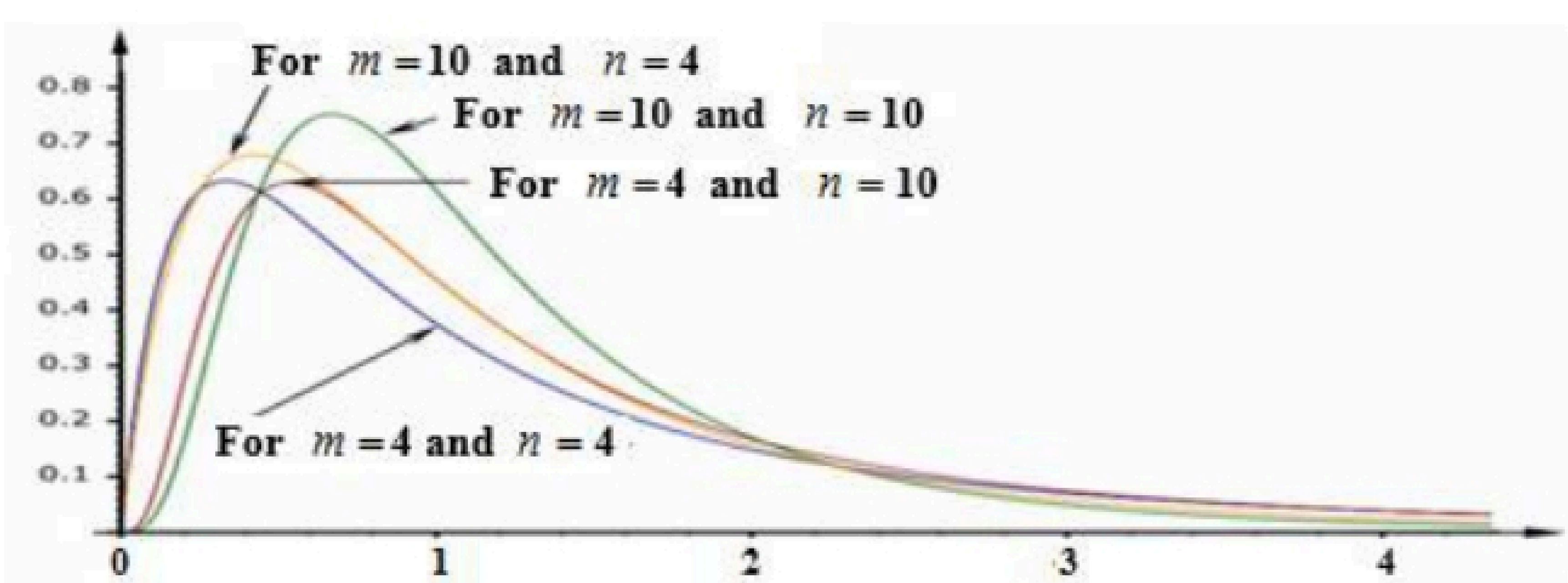
بمَعْلمتين $n$, $m \ni \mathbb{N}$ تُدعيان **درجة حرّية البسط** و**المقام** على الترتيب، إذا كانت دالّة كثافته الاحتمالية مُعطاة من خلال العلاقة الآتية:

$$f_X(x) := \frac{\left(\frac{m}{n}\right)^{\frac{m}{2}}}{\beta\left(\frac{m}{2}, \frac{n}{2}\right)} x^{\frac{m}{2}-1}\left(1+\frac{mx}{n}\right)^{-\frac{m+n}{2}} \mathbf{I}_{(0,\infty)}(x) \qquad ;\forall x \in \mathbb{R} \qquad [6,16\text{-a}]$$

علماً أنَّ $\beta(p,q)$ هي قيمة **دالّة بيتا** عند $p$, $q \ni \mathbb{R}$، أو أن تُكتب بالشكل الآتي أيضاً:

$$f_X(x) = \frac{\Gamma\left(\frac{m+n}{2}\right) m^{\frac{m}{2}} n^{\frac{n}{2}}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \frac{x^{\frac{m}{2}-1}}{(n+mx)^{\frac{m+n}{2}}} \mathbf{I}_{(0,\infty)}(x) \qquad ;\forall x \in \mathbb{R} \qquad [6,16\text{-b}]$$

إنَّ شكل الرسم البياني لدالّة الكثافة الاحتمالية لهذا التوزيع يتغيّر بتغيُّر قيم الوسطاء $m$ و$n$، والشكل الآتي يقدّم لنا عروضاً من أجل قيم مختلفة لكلٍّ من $m$ و$n$.




الشكل (٦,٣١ .أ)

### (٦,٢,٦,٢) ملاحظات

١- سنرمز لدالّة توزيع هذا المتغيِّر العشوائي ب $\mathbf{F}(m;n)$ في حين سنرمز ب $\mathbf{f}(x;m,n)$ لدالّة كثافته الاحتمالية.

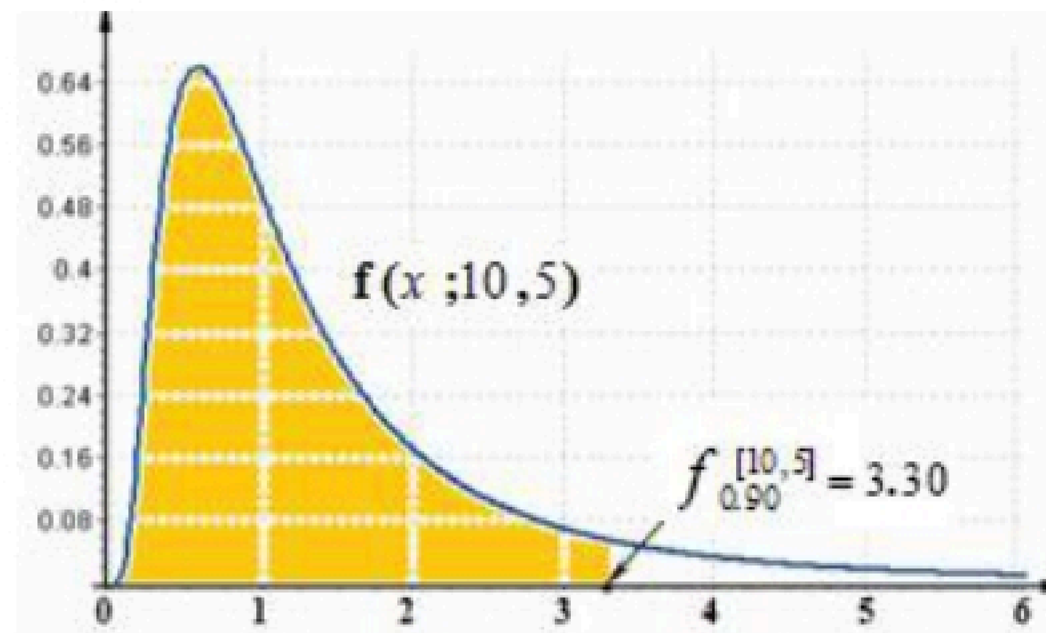
٢- إنَّ هذا النوع من المتغيِّرات العشوائية له استخدامات عديدة منها اختبارات مقارنة التباين للمجتمعات الإحصائية، وفي تحليل التباين، وكذلك في تحليل التغاير (تمام التباين).

٣- سنرمز ب $f_{\alpha}^{[m,n]}$ للقيمة $x$ الواقعة على المحور الأفقي $\mathbf{oX}$ التي يقع على يسارها وتحت منحنى دالّة الكثافة الاحتمالية $\mathbf{f}(x;m,n)$ مساحة قدرها $\alpha$ (قيمة هذه المساحة تساوي $P(X<x)$)، فعلى سبيل المثال من أجل $m=10$ و $n=5$ و $\alpha=0.90$ يكون لدينا $f_{0.90}^{[10,5]}=3.30$، أي إنّه لدينا من أجل هذا التوزيع $P(X<3.30)=0.90$ (للتوضيح انظر (٦,٣١ .ب) الآتي).

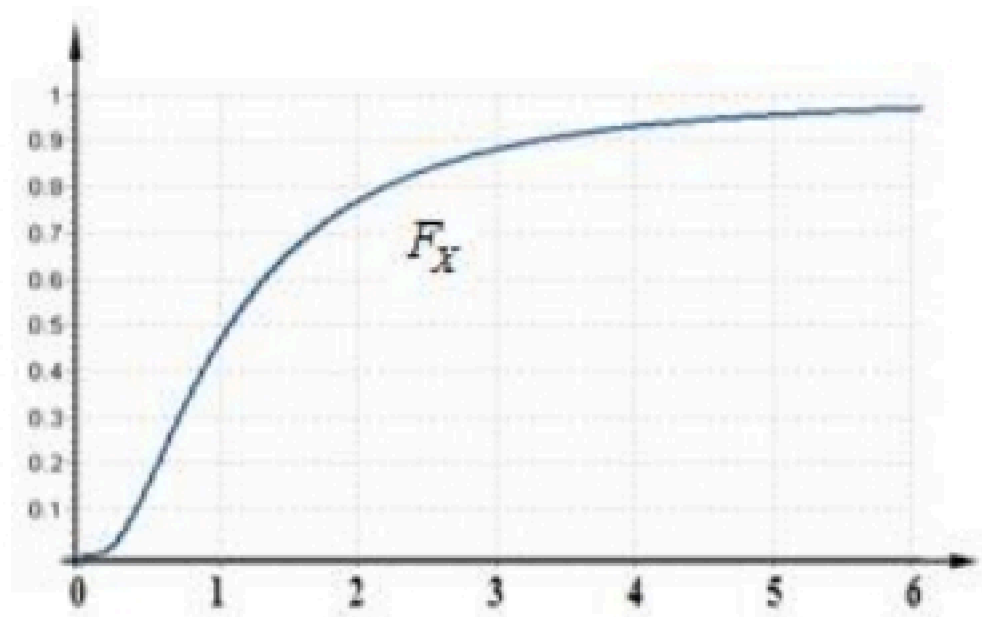
٤- باستخدام تعريف دالّة التوزيع الاحتمالية لمتغيِّر عشوائي مستمر يكون لدالّة توزيع هذا المتغيِّر العشوائي العرض الآتي :

$$F_X(x) = \int_{-\infty}^{x} \frac{\left(\frac{m}{n}\right)^{\frac{m}{2}}}{\beta\left(\frac{m}{2}, \frac{n}{2}\right)} t^{\frac{m}{2}-1}\left(1+\frac{m}{n}t\right)^{-\frac{m+n}{2}} \mathbf{I}_{(0,\infty)}(t)\, dt \qquad ;\forall x \in \mathbb{R}$$

ومن أجل $m = 10$ و $n = 5$ يكون لرسمها البياني الشكل الآتي (٦,٣١.ج)، وأمّا لحساب قيم دالّة التوزيع الاحتمالية السابقة فإنّه يوجد جدول في آخر هذا الكتاب يعطينا قيماً لهذه الدالّة من أجل قيم مختلفة لـ $m$ و $n$، فعلى سبيل المثال: من أجل $m = 8$ و $n = 10$ مع $\alpha = 0.99$ يكون لدينا $f_{0.99}^{[8,10]} = 5.06$، أي إنَّ $P(X < 5.06) = 0.99$ .




الشكل (٦,٣١.ب) المساحة الواقعة على يسار النقطة $x = 3.30$




الشكل (٦,٣١.ج) دالّة التوزيع الاحتمالية لتوزيع فيشر

تطبيقات على هذا التوزيع نجدها لاحقاً في مجال الإحصاء الرياضياتي.

## (٦,٣) تقارب بعض التوزيعات الاحتمالية من بعضها الآخر

**Convergence of some probability distributions to others**

لقد لاحظنا في صيغة الحساب الاحتمالي لمتغيِّر عشوائي خاضع للتوزيع فوق الهندسي أنّه من الصعوبة حساب قيمة الاحتمال عندما تصبح قيم المَعْلمتين $N$ و $M$ كبيرة، ولذلك سنقدِّم فيما يلي بعض المبرهنات (سيترك براهينها كتمارين للقارئ لأنَّها ذات طابع رياضياتي) التي تبيِّن لنا تقارب بعض التوزيعات الاحتمالية من بعضها الآخر لدى تحقُّق شروط مُحدَّدة، ومن ثمَّ يمكننا استخدام توزيع احتمالي عوضاً عن آخر عند تحقُّق هذه الشروط أو على الأقل تحقُّق قيم تسمح لنا باستخدام هذا التقريب؛ وذلك لتجنب الصعوبات في العمليّات الحسابيّة التي قد تواجهنا لدى التوزيع الأصل.

### (٦,٣,١) مبرهنة (بخصوص تقارب التوزيع فوق الهندسي من الحدّاني)

لتكن $N$، $M$ و $n$ من $\mathbb{N}$ مع $N > M$ و $N \ge n$ ، فعندئذٍ من أجل كل $k \ni \mathbb{N}^o$ مع:

$$\max\{0 \, , n + M - N\} \le k \le \min\{n \, , M\}$$

ستكون العلاقة الآتية مُحقَّقة:

$$\lim_{\substack{N \to \infty \\ M \to \infty \\ \frac{M}{N} \to p}} \binom{M}{k} \cdot \binom{N-M}{n-k} \cdot \binom{N}{n}^{-1} = \binom{n}{k} p^k \, (1-p)^{n-k} \qquad [6,17]$$

### (٦,٣,٢,١) نتيجة

من المبرهنة السابقة (٦,٣,١) نستنتج أنّه يمكننا استخدام التوزيع الحدّاني لحساب احتمالات تتعلّق بمتغيِّر عشوائي $X$ خاضع للتوزيع فوق الهندسي، وذلك عندما تكون $N$ و $M$ كبيرة بقدر كافٍ على أنَّ نأخذ $p = \frac{M}{N}$ . إنَّ عمليّة الحساب هذه سينتج عنها خطأ صغير في قيمة الاحتمال، ولكن هذا الخطأ يمكن إهماله عندما يكونُ الاحتمال المطلوب صعب الحساب وفقاً للتوزيع الأصل.

**(٦,٣,٢,٢) ملاحظات**

١- في الجوانب التطبيقية ينصبّ الاهتمام على كبر قيمة $N$ بشكل أساسي، وأن تكون $N >> n$ (أي $N$ أكبر بكثير من $n$). علماً أنَّ مسألة الحكم على القيمة $N$ بأنَّها كبيرة بقدر كاف بالنسبة إلى $n$ تعود إلى إمكانيات المرء وليس هناك معيار مُحدَّد لتبيان قيمة $N$ إن كانت كبيرة بقدر كاف بالنسبة إلى $n$ أم لا، ولكنَّ في الجوانب الإحصائية ينظر إلى أنَّ $N$ كبيرة بقدر كاف بالنسبة إلى $n$ إذا كان $n \geq 0.05N$ التي نجد منها أنَّ $N \leq 20n$ .

٢- إضافة لما سبق فإنَّنا سنصطلح (وعلى سبيل التبسيط) أنَّه إذا كانت قيمة عددية صحيحة $x \geq 100$ (عادةً تُأخذ $x \leq 1000$) فإنَّنا سنتعامل معها على أنَّها كبيرة بقدر كاف (ويُرمز لذلك بـ $x >> 1$)، وكذلك إذا كانت قيمة عددية حقيقية $x \geq 0.01$ (عادةً تُأخذ $x \geq 1000$) فإنَّنا سنتعامل معها على أنَّها صَغيرةٌ بقدر كاف (ويُرمز لذلك بـ $x << 1$).

لقد لاحظنا أنَّ المبرهنة السابقة تسمح لنا باستخدام التوزيع الحداني عوضاً عن التوزيع فوق الهندسي من أجل شروط محدَّدة، ولكن عندما تصبح قيمة $n$ كبيرة بقدر كاف و$p$ صغيرة بقدر كاف أيضاً فإنَّ حساب الاحتمال المطلوب باستخدام التوزيع الحداني يصبح صعباً في كثير من الحالات، ولذلك قُدِّم تقريب آخر للتوزيع الحَداني بحيث يصبح معه حساب الاحتمال المطلوب أكثر يسراً وسهولة.

إنَّ المبرهنة الآتية تسمح لنا باستخدام توزيع بواسوني عوضاً عن التوزيع الحدَّاني عندما تكون $n$ كبيرة بقدر كاف و$p$ صغيرة بقدر كاف.

**(٦,٣,٢) مبرهنة (بخصوص تقارب التوزيع الحدَّاني من البواسوني)**

ليكن $k \in \mathbb{N}$ و $0 < p < 1$ ، فعندئذ من أجل كل $k \in \mathbb{N}^o$ تكون العلاقة الآتية مُحقَّقة:

$$\lim_{\substack{n \to \infty \\ p \to 0 \\ n \cdot p \to \lambda}} \binom{n}{k} p^k \, (1-p)^{n-k} = \frac{\lambda^k}{k\,!} \, \mathrm{e}^{-\lambda} \qquad [6,18]$$

**(٦,٣,٢,١) نتيجة**

نستنتج من المبرهنة السابقة (٦,٣,٢) أنَّه عندما تكون $n >> 1$ وكذلك $p << 1$، فإنَّه يمكننا استخدام توزيع بواسون لحساب احتمالات تتعلَّق بمتغيِّر عشوائي $X$ خاضع للتوزيع الحدَّاني على أن نأخذ قيمة الوسيط $\lambda = n \cdot p$ ، وفي هذه الحالة ينشأ لدينا خطأ صغير يمكن إهماله بسبب صعوبة حساب هذا الاحتمال باستخدام التوزيع الأصل.

تواجهنا في الواقع صعوبات في حساب احتمال يتعلَّق بالتوزيع الحدَّاني عندما تكون $n >> 1$ وكذلك $p << 1$، وهي ذات الصعوبات التي تواجهنا في حساب احتمال يتعلَّق بالتوزيع الحدّاني السالب عندما تكون $n >> 1$ و $p$ قريبة جداً من الواحد، ولهذا السبب قدِّم تقريب لحساب احتمال يتعلَّق بالتوزيع الحدّاني السالب باستخدام التوزيع البواسوني أيضاً، وهذا ما تبيّنه لنا المبرهنة الآتية.

**(٦,٣,٣) مبرهنة (بخصوص تقارب التوزيع الحدَّاني السالب من البواسوني)**

من أجل كل $k \in \mathbb{N}^o$ ، $n \in \mathbb{N}$ و $0 < p < 1$ يكون لدينا ما يلي مُحقَّقاً:

$$\lim_{\substack{n \to \infty \\ p \to 1 \\ \frac{n(1-p)}{p} \to \lambda}} \binom{k+n-1}{k} p^n \, (1-p)^k = \frac{\lambda^k}{k\,!} \, \mathrm{e}^{-\lambda} \qquad [6,19]$$

**(٦,٣,٣,١) نتيجة**

من هذه المبرهنة نستنتج أنَّه يمكننا استخدام التوزيع البواسوني عوضاً عن الحدَّاني السالب في حساب الاحتمال المطلوب على

أن نأخذ $\lambda =: \frac{n\,(1-p)}{p}$ ، وبعد الأخذ بالحسبان أنَّ $n$ كبيرة بقدر كافٍ و $p$ قريبة من الواحد بقدر كافٍ.

**أمثلة (٢, ٣, ٣, ٦)**

١- يوجد في جامعة 10000 طالب منهم 10 طلاب متميّزين على مستوى جامعات البلد، فعندئذ:

أ- إذا كنّا سنخضع كل طالب وبشكل مستقل عن الآخر لاختبار التميّز، فما هو احتمال أن يكون لدينا طالبين متميّزين من بين الطلاب المختبرين في الحالتين الآتيتين:

أ-١- إذا تمَّ سحب عشوائي ل 50 طالباً من طلاب هذه الجامعة؟

أ-٢- إذا تمَّ سحب عشوائي ل 500 طالب من طلاب هذه الجامعة؟

ب- إذا كنّا سنخضع كل طالب للاختبار، وأنّنا سنكتفي بالعدد الذي يوافق حصولنا على أول طالب متميّز، فعندئذ:

ب-١- ما هو احتمال أن نجري خمسة اختبارات فقط؟

ب-٢- ما هو احتمال أن نجري مئة اختبار للحصول على هذا الطالب المتميّز؟

ج- سُحبت عينة عشوائية (دفعةً واحدةً) بحجم 500 طالب، فما هو احتمال أن تحوي هذه العينة على جميع الطلاب المتميّزين؟

د- إذا كنّا سنخضع كل طالب وبشكل مستقل عن الآخر لاختبار التميز، وأنّنا سنتابع في تنفيذ الاختبارات على الطلاب حتى الحصول على ثلاثة طلاب متميِّزين فنتوقف عن الاختبارات، فما هو احتمال أن نقوم بتنفيذ تسعة اختبارات؟

**الحل**: من أجل الطلب:

أ- لناخذ $X$ متغيراً عشوائياً راصداً لعدد الطلاب المتميِّزين الذين سنحصل عليهم نتيجة للاختبارات التي سنجريها، فعندئذ سيكون هذا المتغيِّر العشوائي خاضعاً للتوزيع الحدّاني، أي إنَّ الاحتمال المطلوب يحسب بالعلاقة الاتية:

$$P(X=k)=\binom{n}{k} p^{k}\,(1-p)^{n-k} \qquad k=0,1,2,\ldots,n$$

ومن ثمَّ يكون لدينا من أجل الطلب:

أ-١- لدينا مَعَالم هذا التوزيع هي $n=50$ و $p=\frac{10}{10000}=0.001$ ، والاحتمال المطلوب هو $P(X=2)$ ، فيكون لدينا:

$$P(X=2)=\binom{50}{2} 0.001^{2}\,(1-0.001)^{48}=1225\times10^{-6}\times0.953=0.00117$$

أ-٢- لدينا مَعَالم هذا التوزيع هي $n=500$ و $p=\frac{10}{10000}=0.001$ ، والاحتمال المطلوب هو $P(X=2)$ ، حيث نجد:

$$P(X=2)=\binom{500}{2} 0.001^{2}\,(1-0.001)^{498}=124750\times10^{-6}\times0.6076=0.0757981$$

ونلاحظ هنا أنَّه يمكننا استخدام توزيع بواسون بدلاً عن التوزيع الحدّاني على أن نأخذ قيمة المَعْلَمة $\lambda$ كما يلي:

$$\lambda=n\cdot p=500\times0.001=0.5$$

فيكون لدينا:

$$P(X=2)=\frac{0.5^{2}}{2!}\mathrm{e}^{-0.5}=0.075816$$

فيُلاحظ أنَّ القيمتين قريبتين جداً من بعضهما، ولكن حساب المقدار الأخير أكثر سهولة.

ب- ليكن $X$ متغيراً عشوائياً راصداً لعدد الاختبارات التي سنجريها حتى حصولنا على أول طالب متميّز، فعندئذ سيخضع

هذا المتغيِّر العشوائي إلى التوزيع الهندسي، أي إنَّ الاحتمال المطلوب يحسب بالعلاقة الاتية:

$$P(X=k)=(1-p)^{k-1}\cdot p$$

حيث لدينا $p=0.001$ ، ومن ثَّم يكون من أجل الطلب:

ب-١- الاحتمال المطلوب هو:

$$P(X=5)=(0.999)^4\ 0.001=0.000996$$

ب-٢- الاحتمال المطلوب هو:

$$P(X=100)=(0.999)^{99}\ 0.001=0.0009057$$

وهذا يعني أنَّ احتمال حصولنا على طالب متميِّز باختبارات سابقة سيكون أكبر.

ج- لنأخذ $X$ متغيّراً عشوائياً راصداً لعدد الطلاب المتميّزين في العينة المسحوبة، فعندئذ سيكون لهذا المتغيِّر العشوائي توزيع فوق هندسي، أي إنَّ الاحتمال المطلوب يحسب بالعلاقة الآتية:

$$P(X=k)=\binom{M}{k}\cdot\binom{N-M}{n-k}\cdot\binom{N}{n}^{-1}$$

ومن أجل مسألتنا هذه لدينا $N=10000$ ، $M=10$ و $n=500$ ، ومن ثَّم الاحتمال المطلوب هو:

$$P(X=10)=\binom{500}{10}\cdot\binom{4500}{190}\cdot\binom{10000}{500}^{-1}$$

وبملاحظة صعوبة حساب المقدار الأخير فإنَّنا سنستخدم التوزيع الحداني في حساب الاحتمال المطلوب عوضاً عن التوزيع فوق الهندسي، أي سنستخدم العلاقة الآتية لحساب الاحتمال المطلوب:

$$P(X=k)=\binom{n}{k}p^k\ (1-p)^{n-k}$$

على أن نأخذ قيم المَعَالم $n=500$ و $p=\frac{M}{N}=0.001$ ، ومن ثَّم يكون الاحتمال المطلوب هو:

$$P(X=10)=\binom{500}{10}0.01^{10}\ (0.999)^{990}$$

ويمكننا هنا أن نتابع بتقريب التوزيع السّابق إلى توزيع بواسوني أيضاً، ومن ثَّم نستخدم العلاقة الآتية لحساب الاحتمال المطلوب:

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda}$$

على أن نأخذ قيمة المَعْلمة $\lambda=n\cdot p=500\times0.001=0.5$ ، ولذلك يكون لدينا:

$$P(X=10)=\frac{0.5^{10}}{10!}e^{-0.5}=1.63226\times10^{-10}$$

د- ليكن $X$ متغيّراً عشوائياً راصداً لعدد الاختبارات المنفَّذة التي نفشل فيها حتى الحصول على ثلاثة طلاب متميّزين، فعندئذ سيكون لهذا المتغيِّر العشوائي توزيع حدّاني سالب بمَعْلمتين $n=3$ و $p=0.001$ ، ومن ثَّم سنستخدم العلاقة [6,5-a] لحساب الاحتمال المطلوب، حيث لدينا:

$$P(X=k)=\binom{k+n-1}{k}\cdot p^{n}\cdot(1-p)^{k}$$

ومنه يكون الاحتمال المطلوب هو:

$$P(X=6)=\tilde{\mathbf{b}}(6;3,0.001)=\binom{6+3-1}{6}(0.001)^{3}(0.999)^{6}=2.7832\times10^{-8}$$

بالطبع كان من الممكن أن نوجد الحل بالطريقة الآتية أيضاً.

ليكن $X$ متغيّراً عشوائياً راصداً لعدد الاختبارات المنفّذة حتى الحصول على ثلاثة طلاب متميّزين، فعندئذ سيكون لهذا المتغيِّر العشوائي توزيع حدّاني سالب بمَعْلمتين $n=3$ و $p=0.001$ حيث لدينا العلاقة [6,5-c] :

$$P(X=k)=\binom{k-1}{n-1}p^{n}\ (1-p)^{k-n}$$

التي نجد منها ذات النتيجة التي حصلنا عليها سابقاً، حيث لدينا:

$$P(X=9)=\binom{9-1}{2}(0.001)^{3}(0.999)^{6}=2.7832\times10^{-8}$$

**هذا ما تيسّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل**

# تمارين الفصل السادس

١- بيّن أيّ من الدوال الآتية تصلح لأن تكون دالّة كثافة احتمالية لمتغيّر عشوائي:

$$\textbf{a)}\ f(x)=\begin{cases}\frac{8}{3}x^2 & for\ x\in[-0.5\,,1]\\ 0 & otherwise\end{cases}\quad \& \quad \textbf{b)}\ f(x)=\begin{cases}\frac{1}{3}x & for\ x\in(-1\,,5)\\ 0 & otherwise\end{cases}$$

$$\textbf{c)}\ f(x)=\begin{cases}\frac{1}{8} & for\ x\in[1\,,4)\\ 0 & otherwise\end{cases}\quad \& \quad \textbf{a)}\ f(x)=\begin{cases}\frac{3}{26}x^2 & for\ x\in[1\,,3]\\ 0 & otherwise\end{cases}$$

٢- بيّن أيّ من الدوال الآتية تصلح لأن تكون دالّة توزيع احتمالية لمتغيّر عشوائي:

$$\textbf{a)}\ F(x)=\begin{cases}\frac{1}{2} & for\ x\le -1\\ 2x & for\ -1<x\le +1\\ 1 & otherwise\end{cases}\quad \& \quad \textbf{b)}\ F(x)=\begin{cases}0 & for\ x\le 5\\ x^3 & for\ 5<x\le 10\\ 0.95 & otherwise\end{cases}$$

$$\textbf{b)}\ F(x)=\begin{cases}0 & for\ x\le -5\\ e^x & for\ -5<x\le 0\\ 1 & otherwise\end{cases}\quad \& \quad \textbf{c)}\ F(x)=\begin{cases}e^x & for\ x\le 0\\ x+0.5 & for\ 0<x\le 0.5\\ 1 & otherwise\end{cases}$$

٣- بيّن أيّ مما يلي تصلح لأن تكون دالّة كتلة احتمالية لمتغيّر عشوائي:

$$\textbf{a)}\ P(X=x)=\frac{x}{6}\ \ ;x\in\mathbb{N}_6 \quad \& \quad \textbf{b)}\ P(X=k)=\frac{1}{100}\ \ ;k\in\mathbb{N}_{100}$$

$$\textbf{c)}\ P(X=x)=\frac{2^k}{k\,!}\,\mathrm{e}^{-2}\ \ ;k=0,1,2 \quad \& \quad \textbf{c)}\ P(X=x)=\frac{1}{4}-x\ \ ;x=\frac{1}{2},\frac{1}{4},\frac{1}{6},\frac{1}{8}$$

٤- بيّن أيّ مما يلي تصلح لأن تكون دالّة توزيع احتمالية لمتغيّر عشوائي متقطّع:

$$\textbf{a)}\ F(x)=\begin{cases}\frac{1}{2} & for\ x\le 0\\ 0.5 & for\ 0<x\le 4\\ 1 & otherwise\end{cases}\quad \& \quad \textbf{b)}\ F(x)=\begin{cases}\mathrm{e}^x & for\ x\le 0\\ 1 & otherwise\end{cases}$$

$$\textbf{b)}\ F(x)=\begin{cases}0 & for\ x\le -2\\ 0.75 & for\ -2<x\le 10\\ 1 & otherwise\end{cases}\quad \& \quad \textbf{c)}\ F(x)=\begin{cases}0 & for\ x\le 70\\ 0.25 & for\ 70<x\le 105\\ 0.99 & otherwise\end{cases}$$

٥- يوجد في مدينة 4000000 سيّارة منها 4000 سيارة لم تقم بتجديد رخصة السير، فقامت دورية شرطة المرور بإنشاء حاجز على طريق ما في هذه المدينة للتحقّق من صلاحية رخصة السير. عندئذ:

أ- إذا تمّ استجواب 500 سائق سيارة على هذا الطريق، فما هو احتمال أن يكون قد تمّ ضبط 10 سيارات تسير برخصة سير منتهية الصلاحية؟

ب- إذا كان الاستجواب سوف يستمر حتى يتمّ ضبط أول سيّارة تسير برخصة سير منتهية الصلاحية فتوقفها وتذهب بها إلى

الحجز وتنهي عمليّة التفتيش، فما هو احتمال أن يكون قد تمَّ استجواب سائقي 50 سيارة؟

ج- إذا كان الاستجواب سوف يستمر حتى يتمَّ ضبط 10 سيّارات تسير برخصة سير منتهية الصلاحية فتأخذ هذه السيارات إلى الحجز وتنهي عمليّة التفتيش، فما هو احتمال أن يكون قد تمَّ استجواب سائقي 100 سيارة حتى إنهاء عملية التفتيش؟

د- إذا قامت الدورية بعملية تفتيش لجميع السيارات الموجودة في حي معيَّن والحاوي على 5000 سيارة، فما هو احتمال أن يكون قد تمَّ ضبط 30 سيّارة برخصة سير منتهية الصلاحية؟

٦- يقوم رجل بالرمي على هدف، فإذا علمنا أنَّ احتمال إصابة الهدف من قبل الرجل يساوي 0.85، وأنَّ الرجل بعد كل رمية تُؤخذ منه البندقة ويُعطى بندقية أخرى مختلفة عن سابقتها، وكذلك لا يمكنه معرفة نتيجة رميته على الهدف (أي الرميات على الهدف يتم بشكل مستقل كل منها عن الآخر)، فعندئذ:

أ- كم رمية يجب على الرامي أنَّ يطلقها على الهدف حتى يصبح احتمال إصابة الهدف أكبر من 0.9999؟

ب- بفرض أنَّ الرجل أطلق 12 رمية على الهدف فما هو احتمال أنَّ يصيب بسبع منها؟

ج- بفرض أنَّه سيُطلب من الرجل التوقف عن الرمي (دون علمه بذلك مسبقاً) لدى تحقيقه إصابه للهدف، فما هو احتمال أنَّ يكون قد أطلق الرجل خمس رميات على الهدف فقط؟

د - بفرض أنَّه سيُطلب من الرجل التوقف عن الرمي (دون علمه بذلك مسبقاً) لدى تحقيقه ثلاث إصابات للهدف، فما هو احتمال أنَّ يكون قد أطلق الرجل سبع رميات على الهدف فقط؟

٧- ليكن $X$ متغيِّراً عشوائياً قانون توزيعه مُعطى من خلال الجدول الآتي:

| $i$ | 1 | 2 | 3 | 4 | sum |
|---|---|---|---|---|---|
| قيم $X$ $x_i$ | 0 | 1 | 2 | 3 | |
| $p_i = P(X = x_i)$ | $\frac{16}{64}$ | $\frac{16}{64}$ | $\frac{16}{64}$ | $\frac{16}{64}$ | 1 |

والمطلوب تعيين نوع هذا المتغيّر العشوائي، ومن ثَّم حساب الاحتمالات الآتية:

**a)** $P(X \geq 1) = ?$ **b)** $P(0.5 \leq X < 2.5) = ?$ **c)** $P(X \leq 1.25) = ?$

٨- ليكن $X$ متغيِّراً عشوائياً قانون توزيعه مُعطى من خلال الجدول الآتي:

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | sum |
|---|---|---|---|---|---|---|---|
| قيم $X$ $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 | |
| $p_i = P(X = x_i)$ | 0.04 | 0.20 | 0.30 | 0.40 | ? | 0.05 | |

علماً أنَّ **?** تشير إلى أنَّ قيمة الاحتمال مجهولة، والمطلوب ما يلي:

أ- تعيين قيمة الاحتمال المجهولة.

ب- تعيين نوع هذا المتغيّر العشوائي، ومن ثَّم حساب الاحتمالات الآتية:

**a)** $P(X < 4) = ?$ **b)** $P(X < 0.5) = ?$ **c)** $P(X > 1.25) = ?$

٩- في نظام الاتحاد الرياضي لبلد ما أنَّ الفريق الذي يفوز ب 10 مباريات من أصل 12 مباراة يمكنه الانتقال من أندية الدرجة الثانية إلى أندية الدرجة الأولى. فلو افترضنا أن كل مباراة تجري بشكل مستقل عن الأخرى، فعندئذ:

أ- إذا كان احتمال فوز فريق A في أية مباراة يساوي 0.65 وأنَّه قد فاز بأربع مباريات من أصل ست مباريات، فما هو احتمال أن يحقِّق الفريق A الانتقال إلى أندية الدرجة الأولى؟

ب- إذا كان احتمال فوز فريق A في أية مباراة يساوي 0.65 وأنَّه قد فاز بأربع مباريات من أصل ست مباريات، وهناك فريق آخر B احتمال فوزه في أية مباراة يساوي 0.60 وأنَّه قد فاز بست مباريات من أصل ثمان مباريات، فأي الفريقين له احتمال أكبر في الانتقال إلى أندية الدرجة الأولى؟

١٠- إذا علمت أنَّ احتمال أن تنجح عمليّة جراحيّة من نوع معيَّن لمريض هو 0.85، وأنَّ العمليّات تُجرى بشكل مستقل كل منها عن الأخرى، فعندئذ:

أ- ما هو احتمال نجاح العمليّة لخمسة مرضى من أصل 7 مرضى ستجرى لهم هذه العملية؟

ب- إذا كان طاقم العمل الجراحي سيستمر في إجراء هذا النوع من العمليّات الجراحيّة على المرضى حتى حدوث أول فشل لهذه العمليّة حيث يتوقف الطاقم عن تنفيذ هذا النوع من العمليّات، فما هو احتمال أن يقوم هذا الطاقم بتنفيذ 15 عمليّة جراحيّة؟

ج- إذا كان طاقم العمل الجراحي سيستمر في إجراء هذا النوع من العمليّات الجراحيّة على المرضى حتى يتبيَّن لهم فشلها للمرَّة الثّالثة حيث يتوقف الطاقم عن تنفيذ هذا النوع من العمليّات الجراحيّة، فما هو احتمال أن يقوم هذا الطاقم بتنفيذ 30 عمليّة جراحيّة قبل توقفه عن تنفيذ هذا النوع من العمليّات؟

١١- لدى شركة طيران مدنية نوعين من الطائرات تعمل بالنوع نفسه من المحركات، ولكن النوع الأول يعمل بأربع محركات، وأمّا النوع الثاني فيعمل بمحركين فقط. لو افترضنا أنَّ محركات الطائرات تعمل بشكل مستقل كل منها عن الأخرى وفي كلٍّ من النوعين، وأنَّ احتمال تعطّل محرك عن العمل أثناء الطيران يساوي 0.001، وأنَّ الطائرة يمكن لها أن تتمّ رحلتها إذا كان نصف عدد محركاتها على الأقل في حالة العمل. عندئذ بيّن أي النوعين من الطائرات أكثر أماناً للوصول إلى هدفها.

١٢- في إحدى الكليات يوجد 120 محاضراً (من مختلف المراتب العلمية) و30 قائماً بالأعمال. أرادت عمادة الكلية تشكيل لجنة استشارية مكوَّنة من 15 شخصاً يسحبون عشوائياً من الكادر العامل في هذه الكلية، فما هو احتمال أن يكون في هذه اللجنة خمسة محاضرين على الأقل؟

١٣- يشير استطلاع للرأي حول التدخين أنَّه من كلّ 1000 شخص يوجد بينهم 50 شخصاً مدخناً، فلو قمنا بسؤال 250 شخصاً اختيروا بشكل عشوائي، فما هو احتمال أن يكون من بينهم تسعة أشخاص مدخنين على الأكثر؟

١٤- ليكن $X$ متغيِّراً عشوائياً خاضعاً للتوزيع الطبيعي بمَعْلمتين $\mu = 5$ و$\sigma = 1.5$، والمطلوب حساب الاحتمالات الآتية:

**a)** $P(X \le 7) = ?$ **b)** $P(3 \le X \le 7) = ?$ **c)** $P(X > 8) = ?$ **d)** $P(-5 < X < -2) = ?$

١٥- في مصنع لإنتاج المكانس الكهربائية يوجد أربعة خطوط للإنتاج $L_1$ ، $L_2$ ، $L_3$ و $L_4$ لها القدرة نفسها في الإنتاج، ولكن نسبة المعيب في إنتاج هذه الخطوط هو 0.01، 0.02، 0.03 و0.04 على الترتيب. عندئذ أجب عما يلي:

أ- نقوم بسحب عشوائي لمكنسة من الإنتاج الكلي للمصنع، فما هو احتمال أن تكون المحوِّلة المسحوبة صالحة للاستخدام؟

ب- قمنا بسحب عشرة مكانس واحدةً تلو الأخرى من الإنتاج الكلي للمصنع ومن ثمَّ فحصها، فما هو احتمال حصولنا على مكنستين معيبتين فقط؟

ج- بفرض أن المصنع أنتج 1000 مكنسة وقمنا ببيع 200 مكنسة سحبت عشوائيًّا من هذا الإنتاج، فما هو احتمال أن تكون جميع المكانس التي تمَّ بيعها ليست معيبة؟

د - بفرض أنَّ عمر أي مكنسة ( مقدَّرة بالسنة ) يخضع للتوزيع الأسي بمتوسط ثلاث سنوات، فعندئذ:

١- عيّن قيمة المَعْلمة لتوزيع (للتوزيع الأسيّ) أعمار المكانس؟

٢- أحسب احتمال ألاَّ تُعمِّر مكنسة من إنتاج هذا المصنع أكثر من خمس سنوات؟

٣- إذا عملت هذه المكنسة لثلاث سنوات فما هو احتمال أن تعمل لثلاث سنوات قادمة أيضاً؟

١٦- لنأخذ تجربة إلقاء حجر نرد متوازن لمرّتين متتاليتين، وليكن $X$ متغيِّرًا عشوائيًّا فوق الفضاء الاحتمالي لهذه التجربة معرفاً كما يلي:

$$X(\omega) := \begin{cases} -1 & \textit{if } \omega \textit{ is an odd number} \\ +1 & \textit{if } \omega \textit{ is an even number} \end{cases}$$

عندئذ المطلوب ما يلي :

أ- لأي توزيع يخضع هذا المتغير العشوائي، ولماذا؟

ب- عيِّن دالَّة توزيع هذا المتغيِّر العشوائي، ومن ثمَّ ارسمها.

ج- أستخدم دالَّة التوزيع لحساب الاحتمالات الآتية:

**a)** $P(X \geq -0.5) = ?$ **b)** $P(-1 < X \leq 0.5) = ?$ **c)** $P(X < 0.8) = ?$

## الفصل السابع

# استقلال وعزوم المتغيرات العشوائية

# INDEPENDENCE AND MOMENTS OF RANDOM VARIABLES

## (٧,١) استقلال المتغيرات العشوائية

**Independence of Random Variables**

لقد قدَّمنا في الفصل الرابع دراسة الاستقلال للحوادث، وفيما يلي سنقدّم دراسة مبسطة وموجزة حول استقلال المتغيِّرات العشوائية.

### (٧,١,١) الاستقلال لعدد منته من المتغيِّرات العشوائية

نقدّم في هذه الفقرة مفهوم الاستقلال لعدد منته من المتغيِّرات العشوائية مع بعض الاختبارات التي تساعدنا في اتخاذ قرار بشأن ذلك.

#### (٧,١,١,١) تعريف (الاستقلال لعدد منته من المتغيِّرات العشوائية)

لتكن $X_1$، $X_2$، ...و $X_n$ متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، علماً أنَّ $n \geq 2$ عدد صحيح، فعندئذ يُقال عن هذه المتغيِّرات العشوائية إنَّها **مستقلّة عشوائياً** (**أو مستقلّة إحصائيا**) إذا كان من أجل أيّة حوادث (**مجموعات بوريلية**) $A_1$، $A_2$، ...و $A_n$ من $\Re$ لدينا العلاقة الآتية مُحقَّقة:

$$P(X_1 \in A_1, X_2 \in A_2, ..., X_n \in A_n) = \prod_{k=1}^{n} P(X_k \in A_k) \qquad [7,1]$$

علماً أنَّ $\prod$ يرمز إلى الضرب (**المعروف**) بين المقادير العدديّة.

#### (٧,١,١,٢) ملاحظات

١- حيثما ترد عبارة متغيِّرات عشوائية **مستقلّة** إنَّما يُقصد بها متغيِّرات عشوائية **مستقلّة عشوائياً**.

٢- في الحالة الخاصّة، إذا كان من أجل كلّ $i$ و $j$ من $\mathbb{N}_n$ مع $i \neq j$ لدينا:

$$P(X_i \in A_i, X_j \in A_j) = P(X_i \in A_i) \cdot P(X_j \in A_j) \qquad [7,2]$$

فعندئذ يُقال عن المتغيِّرات العشوائية $X_1$، $X_2$، ...و $X_n$ إنَّها **مستقلّة مثنى مثنى**.

#### (٧,١,١,٣) مثال

ليكن $X$، $Y$ و $Z$ متغيِّرات عشوائية فوق الفضاء الاحتمالي لتجربة إلقاء قطعة نقود متوازنة لمرّتين متتاليتين، ومعرَّفةً من أجل كلّ $\omega \in \Omega$ كما يلي:

$$X(\omega)=\mathbf{I}_{\{HH,HT\}}(\omega) \quad \& \quad Y(\omega)=\mathbf{I}_{\{HH,TH\}}(\omega) \quad \& \quad Z(\omega)=\mathbf{I}_{\{HH,TT\}}(\omega)$$

فنجد أنَّ:

$$P(X=1)=P(X=0)=P(Y=1)=P(Y=0)=P(Z=1)=P(Z=0)=\frac{1}{4}$$

$$P(X=1,Y=1)=P(X=0,Y=1)=P(X=1,Y=0)=P(X=0,Y=0)=\frac{1}{4}$$

$$P(X=1,Z=1)=P(X=0,Z=1)=P(X=1,Z=0)=P(X=0,Z=0)=\frac{1}{4}$$

$$P(Y=1,Z=1)=P(Y=0,Z=1)=P(Y=1,Z=0)=P(Y=0,Z=0)=\frac{1}{4}$$

ومن هذه العلاقات ينتج لدينا أنَّ المتغيِّرات العشوائية $X$، $Y$ و$Z$ مستقلّة مثنى مثنى، ولكنَّها ليست مستقلّة وذلك لأنَّ:

$$P(X=1,Y=1,Z=1)=P(\{HH\})=\frac{1}{4} \qquad \& \qquad P(X=1)\cdot P(Y=1)\cdot P(Z=1)=\frac{1}{8}$$

ومن ثمَّ يكون:

$$P(X=1,Y=1,Z=1)\neq P(X=1)\cdot P(Y=1)\cdot P(Z=1)$$

### (٧, ١, ١, ٤) ملاحظة

يوضِّح لنا المثال السّابق أنَّه إذا كانت $X_1$، $X_2$، ...، و $X_n$ مستقلّة مثنى مثنى فليس بالضرورة أنَّ تكون هذه المتغيِّرات العشوائية مستقلّة، ولكن العكس صحيح، وهذا ما ستبيِّنه لنا المبرهنة الآتية.

### (٧, ١, ١, ٥) مبرهنة

لتكن $X_1$، $X_2$، ...، و $X_n$ متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فإذا كانت هذه المتغيِّرات العشوائية مستقلّة، فعندئذ ستكون أيّة أسرة جزئية منها مستقلّة أيضاً.

**البرهان:** لنأخذ $X_{\alpha_1}$، $X_{\alpha_2}$، ...، و $X_{\alpha_s}$ أسرة جزئية من المتغيِّرات العشوائية المُعطاة، علماً أنَّ $1<s\leq n$ وكذلك $\alpha_1\leq n$، $\alpha_2$، ...، و $1\leq\alpha_s$ أعداد صحيحة مختلف بعضها عن بعض، ولنرمز لبقية المتغيِّرات العشوائية بـ $X_{\beta_1}$، $X_{\beta_2}$، ...، و $X_{\beta_r}$ مع $\beta_1\leq n$، $\beta_2$، ...، و $1\leq\beta_r$ أعداد صحيحة مختلف بعضها عن بعض، وكذلك مختلفة عن القيم $\alpha_1$، $\alpha_2$، ...، و $\alpha_s$، علماً أنَّ $r=n-s$، فعندئذ بأخذ $B_1$، $B_2$، ...، و $B_s$ من $\Re$، فإنَّه بسبب استقلال $X_1$، $X_2$، ...، و $X_n$ يمكننا أن نكتب الآتي:

$$\begin{aligned}
P(X_{\alpha_1}\in B_1,X_{\alpha_2}\in B_2,\dots,X_{\alpha_s}\in B_s)&=P\begin{pmatrix}X_{\alpha_1}\in B_1,X_{\alpha_2}\in B_2,\dots,X_{\alpha_s}\in B_s,\\ X_{\beta_1}\in\mathbb{R},X_{\beta_2}\in\mathbb{R},\dots,\in X_{\beta_r}\in\mathbb{R}\end{pmatrix}\\
&=P\left(X_{\alpha_1}\in B_1\right)\cdot P\left(X_{\alpha_2}\in B_2\right)\cdot\ldots\cdot P\left(X_{\alpha_s}\in B_s\right)\cdot\underbrace{P\left(X_{\beta_1}\in\mathbb{R}\right)}_{=1}\cdot\underbrace{P\left(X_{\beta_2}\in\mathbb{R}\right)}_{=1}\cdot\ldots\cdot\underbrace{P\left(X_{\beta_r}\in\mathbb{R}\right)}_{=1}\\
&=P\left(X_{\alpha_1}\in B_1\right)\cdot P\left(X_{\alpha_2}\in B_2\right)\cdot\ldots\cdot P\left(X_{\alpha_s}\in B_s\right)
\end{aligned}$$

وهذا يعني أنَّ المتغيِّرات العشوائية $X_{\alpha_1}$، $X_{\alpha_2}$، ...، و $X_{\alpha_s}$ مستقلّة، وبهذا يتمّ البرهان.

المبرهنة الآتية تعرض لنا إحدى الخصائص المهمَّة المتعلّقة باستقلال المتغيِّرات العشوائية.

**(٦ , ١ , ١ , ٧) مبرهنة**

لتكن $X_1$ ، $X_2$ ،...و $X_n$ متغيِّرات عشوائية مستقلّة فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، ولتكن $g_1$ ، $g_2$ ،... و $g_n$ دوال حقيقية معرَّفة على $\mathbb{R}$ وقيوسة بالنسبة إلى $\Re$ أيضاً، فعندئذ تكون المتغيِّرات العشوائية $g_1(X_1)$ ، $g_2(X_2)$ ،... و $g_n(X_n)$ مستقلّة أيضاً.

**البرهان:** لنأخذ $B_1$ ، $B_2$ ،...و $B_n$ من $\Re$ ، فعندئذ يمكننا أن نكتب الآتي:

$$P\left(g_1(X_1)\in B_1, g_2(X_2)\in B_2,..., g_n(X_n)\in B_n\right)=P\left(X_1\in g_1^{-1}(B_1), X_2\in g_2^{-1}(B_2),..., X_n\in g_n^{-1}(B_n)\right)$$

ومن كون الدوال $g_k$ قيوسة بالنسبة إلى $\Re$ من أجل كلّ $k \in \mathbb{N}_n$ ، فإنَّه سيكون لدينا:

$$g_k^{-1}(B_k)\in\Re \qquad ;\forall k\in\mathbb{N}_n$$

وبما أنَّ $X_1$ ، $X_2$ ،...و $X_n$ مستقلّة، فإنَّه سينتج لدينا ما يلي:

$$\begin{aligned}
&P\left(X_1\in g_1^{-1}(B_1)\, ,\, X_2\in g_2^{-1}(B_2)\, ,..., X_n\in g_n^{-1}(B_n)\right)\\
&\qquad = P\left(X_1\in g_1^{-1}(B_1)\right)\cdot P\left(X_2\in g_2^{-1}(B_2)\right)\cdot ...\cdot P\left(X_n\in g_n^{-1}(B_n)\right)\\
&\qquad = P\left(g_1(X_1)\in B_1\right)\cdot P\left(g_2(X_2)\in B_2\right)\cdot ...\cdot P\left(g_n(X_n)\in B_n\right)
\end{aligned}$$

وهذا يعني أنَّ المتغيِّرات العشوائية $g_1(X_1)$ ، $g_2(X_2)$ ،... و $g_n(X_n)$ مستقلّة، وبهذا يتمّ البرهان.

**(٧ , ١ , ١ , ٧) أمثلة**

١- لتكن $X_1$ ، $X_2$ ،...و $X_n$ متغيِّرات عشوائية مستقلّة فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، ولنأخذ $g_1$ ، $g_2$ ،... و $g_n$ دوال حقيقية معرَّفة على $\mathbb{R}$ كما يلي:

$$g_k(x)=x\pm a_k \qquad ;\forall k\in\mathbb{N}_n\, , x\in\mathbb{R}$$

علماً أنَّ $a_k \in \mathbb{R}$ من أجل كلّ $k \in \mathbb{N}_n$ . عندئذ تكون المتغيِّرات العشوائية الآتية مستقلّة أيضاً:

$$Y_k = g_k(X_k)=X_k\pm a_k \qquad ;\forall k\in\mathbb{N}_n$$

وذلك لأنَّ الدوال $g_k$ مستمرَّة على $\mathbb{R}$ من أجل كلّ $k \in \mathbb{N}_n$ ، ومن ثمَّ ستكون قيوسة بالنسبة إلى $\Re$ أيضاً (**انظر الملحق B**)، ومن ثَمّ بحسب المبرهنة الأخيرة نكون قد أثبتنا صحة ادعائنا.

من التطبيقات المميَّزة لهذا المثال هو أنَّه إذا كانت المتغيِّرات العشوائية $X_1$ ، $X_2$ ،...و $X_n$ مستقلّة وتملك عزوماً ابتدائية من المرتبة الأولى، فعندئذ المتغيِّرات العشوائية:

$$Y_1=X_1-\mathbf{E}X_1 \quad \& \quad Y_2=X_2-\mathbf{E}X_2 \quad \&...\& \quad Y_n=X_n-\mathbf{E}X_n$$

ستكون مستقلّة أيضاً، علماً أنَّ $\mathbf{E}X_i$ هو التوقع الرياضياتي (**أو القيمة الوسطى**) لـ $X_i$ (**للاطلاع على مفهوم التوقع الرياضياتي انظر فقرة العزوم في هذا الفصل**).

٢- لتكن $X_1$ ، $X_2$ ،...و $X_n$ متغيِّرات عشوائية مستقلّة فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، ولنأخذ $g_1$ ، $g_2$ ،... و $g_n$ دوال حقيقية معرَّفة على $\mathbb{R}$ كما يلي:

$$g_k(x)=a_k\, x^2 \qquad ;\forall k\in\mathbb{N}_n\, , x\in\mathbb{R}$$

علماً أنَّ $a_k \in \mathbb{R}^*$ من أجل كلّ $k \in \mathbb{N}_n$ ، فعندئذ نلاحظ أنَّ هذه الدوال مستمرَّة على $\mathbb{R}$ ، ومن ثمَّ تكون قيوسة بالنسبة إلى $\Re$ أيضاً، ومنه بحسب المبرهنة الأخيرة ستكون المتغيِّرات العشوائية الآتية مستقلّة أيضاً:

$$Y_k = g_k(X_k) = a_k X_k^2 \qquad ;\forall k \in \mathbb{N}_n$$

من التطبيقات المهمّة لهذا المثال هو أنّه إذا كانت المتغيِّرات العشوائية $X_1$، $X_2$،...و $X_n$ مستقلّة فإنَّ المتغيِّرات العشوائية $Y_k = X_k^2$ ستكون مستقلّة من أجل كلّ $k \in \mathbb{N}_n$ أيضاً، وإذا كانت المتغيِّرات العشوائية $X_1$، $X_2$،...و $X_n$ تخضع للتوزيع الطبيعي $\mathbf{N}(0,1)$، فعندئذ المتغيِّر العشوائي $Z := \sum_{i=1}^{n} X_i^2$ سيكون خاضعاً لتوزيع كاي مربّع $\chi^2(n)$ (انظر [65]).

٣- لتكن $X_1$، $X_2$،...و $X_n$ متغيِّرات عشوائية مستقلّة، ولنأخذ من أجل كلّ $k \in \mathbb{N}_n$ دوال حقيقية $g_k(x) = \frac{x}{a_k}$ معرَّفة على $\mathbb{R}$ علماً أنَّ $a_k \in \mathbb{R}^*$ من أجل كلّ $k \in \mathbb{N}_n$، فعندئذ نلاحظ أنَّ الدوال $g_k(x) = \frac{x}{a_k}$ مستمرّة على $\mathbb{R}$، ومن ثمّ تكون قيوسة بالنسبة إلى $\Re$، وبحسب المبرهنة الأخيرة تكون المتغيِّرات العشوائية الآتية مستقلّة أيضاً:

$$Y_k = g_k(X_k) = \frac{X_k}{a_k} \qquad ;\forall k \in \mathbb{N}_n$$

من التطبيقات المهمّة لهذا المثال هو أنّه إذا كان $X_1$ و $X_2$ متغيِّرين عشوائيين مستقلّين، فإنَّ المتغيِّرين العشوائيين $Y_1 = X_1$ و $Y_2 = \frac{X_2}{n}$ سيكونان مستقلّين أيضاً، علماً أنَّ $n \in \mathbb{N}$ عدد مثبَّت، وإذا كان $X_1$ خاضعاً للتوزيع $\mathbf{N}(0,1)$ و $X_2$ خاضعاً لتوزيع $\chi^2(n)$، فعندئذ يُبرهن على أنَّ المتغيِّر العشوائي:

$$Z := \frac{Y_1}{\sqrt{Y_2}} = \frac{X_1}{\sqrt{X_2}}\sqrt{n}$$

سيكون خاضعاً لتوزيع ستودنت $\mathbf{T}(n)$ (انظر [65])، ومن أجل $n = 1$ سيصبح لـ $Z$ توزيع كوشي $\mathbf{C}(0,1)$.

في الحقيقة يمكن الاستفادة من التوزيعات الهامشية في تبيان إن كانت مجموعة منتهية من المتغيِّرات العشوائية مستقلّة أم لا، وذلك من خلال استخدام المتَّجهات العشوائيَّة كنقطة انطلاق للدراسة. إنَّ هذه الدراسة تقدِّمها لنا المبرهنات الآتية التي سنبرهن على إحداها ونقبل الأخريات دون برهان.

**مبرهنة (٧, ١, ١, ٨)**

ليكن $\mathbb{X}_n = (X_1, X_2, \dots, X_n)$ متَّجهاً عشوائياً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ بدالَّة توزيع $F_{\mathbb{X}_n}$، فعندئذ تكون المتغيِّرات العشوائيَّة $X_1$، $X_2$،...و $X_n$ مستقلّة **إذا وفقط إذا** كان من أجل كلّ $\vec{x}^n = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ العلاقة الآتية مُحقَّقة:

$$F_{\mathbb{X}_n}(\vec{x}^n) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \dots \cdot F_{X_n}(x_n) \qquad [7,3]$$

علماً أنَّ $F_{X_1}$، $F_{X_2}$،...و $F_{X_n}$ هي دوال التوزيع الهامشيّة للمتغيِّرات العشوائيَّة $X_1$، $X_2$،...و $X_n$ على الترتيب.

**البرهان:**

**من أجل لزوم الشرط:** لدينا المتغيِّرات العشوائيَّة $X_1$، $X_2$،...و $X_n$ مستقلّة، فعندئذ من أجل كلّ $\vec{x}^n \in \mathbb{R}^n$ يمكننا أن نكتب الآتي:

$$F_{\mathbb{X}_n}(\vec{x}^n) = P\left(\mathbb{X}_n \in \underset{k=1}{\overset{n}{\times}}(-\infty, x_k)\right) = P(X_1 < x_1, X_2 < x_2, \dots, X_n < x_n)$$
$$= P(X_1 < x_1) \cdot P(X_2 < x_2) \cdot \dots \cdot P(X_n < x_n) = F_{X_1}(x_1) \cdot F_{X_2}(x_2) \cdot \dots \cdot F_{X_n}(x_n)$$

علماً أنَّ:

$$\underset{k=1}{\overset{n}{\times}}(-\infty, x_k) := (-\infty, x_1) \times (-\infty, x_2) \times \dots \times (-\infty, x_n) \in \Re^n$$

**من أجل كفاية الشرط:** يمكننا أن نكتب الآتي من أجل كلّ $\vec{x}^n \in \mathbb{R}^n$:

$$F_{\mathbb{X}_n}(\vec{x}^n) = P\left(\mathbb{X}_n \in \underset{k=1}{\overset{n}{\times}}(-\infty, x_k)\right)$$
$$= P(X_1 \in (-\infty, x_1), X_2 \in (-\infty, x_2), \ldots, X_n \in (-\infty, x_n)) \qquad \textbf{(1)}$$
$$= P(X_1 \in (-\infty, x_1)) \cdot P(X_2 \in (-\infty, x_2)) \cdot \ldots \cdot P(X_n \in (-\infty, x_n))$$

ومن جهة أخرى لدينا من أجل كلّ $\vec{x}^n \in \mathbb{R}^n$ ما يلي مُحقَّقاً أيضاً:

$$F_{X_1}(x_1) \cdot F_{X_2}(x_1) \cdot \ldots \cdot F_{X_n}(x_n) = P(X_1 < x_1) \cdot P(X_2 < x_2) \cdot \ldots \cdot P(X_n < x_n)$$
$$= P(X_1 \in (-\infty, x_1)) \cdot P(X_2 \in (-\infty, x_2)) \cdot \ldots \cdot P(X_n \in (-\infty, x_n)) \qquad \textbf{(2)}$$

ومن ثمَّ ينتج من العلاقتين **(1)** و**(2)** واستخدام الفرض (تحقُّق العلاقة [7,3]) أنَّه من أجل كلّ $\vec{x}^n \in \mathbb{R}^n$ ستكون العلاقة الآتية مُحقَّقة أيضاً:

$$P\left(\mathbb{X}_n \in \underset{k=1}{\overset{n}{\times}}(-\infty, x_k)\right) = P(X_1 \in (-\infty, x_1)) \cdot P(X_2 \in (-\infty, x_2)) \cdot \ldots \cdot P(X_n \in (-\infty, x_n))$$

وهذا يعني بحسب تعريف الاستقلال أنَّ المتغيِّرات العشوائيَّة $X_1$، $X_2$، ...و $X_n$ مستقلّة.

**(٧, ١, ١, ٩) مبرهنة (تقبل دون برهان)**

ليكن $\mathbb{X}_n = (X_1, X_2, \ldots, X_n)$ متَّجهاً عشوائياً متقطّعاً، فعندئذ تحقُّق الشرط الآتي:

$$p_{i_1 i_2 \ldots i_n} = p_{i_1 \bullet \ldots \bullet} \cdot p_{\bullet i_2 \bullet \ldots \bullet} \cdot \ldots \cdot p_{\bullet \ldots \bullet i_n} \qquad [7,4]$$

هو **شرط لازم وكاف** لاستقلال المتغيِّرات العشوائيَّة $X_1$، $X_2$، ...و $X_n$ المكوِّنة لمركِّبات هذا المتَّجه العشوائي.

**(٧, ١, ١, ١٠) مبرهنة (تقبل دون برهان)**

لتكن $X_1$، $X_2$، ...و $X_n$ متغيِّرات عشوائيَّة متقطِّعة مجموعة قيمها على الترتيب هي:

$$\mathbf{X}_1 = \{x_{i_1} ; i_1 \in I_1\} \quad \& \quad \mathbf{X}_2 = \{x_{i_2} ; i_2 \in I_2\} \quad \& \ldots \& \quad \mathbf{X}_n = \{x_{i_n} ; i_n \in I_n\}$$

وبفرض أنَّ:

$$p_{i_1} = P(X = x_{i_1}) \quad ; i_1 \in I_1 \quad \& \; p_{i_2} = P(X = x_{i_2}) \quad ; i_2 \in I_2 \quad \& \ldots \& \; p_{i_n} = P(X = x_{i_n}) \quad ; i_n \in I_n$$

وكذلك:

$$p_{i_1 i_2 \ldots i_n} = P(X = x_{i_1}, X = x_{i_2}, \ldots, X = x_{i_n})$$

فعندئذ تحقُّق العلاقة الآتية من أجل كلّ $i_1 \in I_1$، $i_2 \in I_2$، ...و $i_n \in I_n$:

$$p_{i_1 i_2 \ldots i_n} = p_{i_1} \cdot p_{i_2} \cdot \ldots \cdot p_{i_n} \qquad [7,5]$$

هو **شرط لازم وكاف** لاستقلال المتغيِّرات العشوائيَّة $X_1$، $X_2$، ...و $X_n$.

**(٧, ١, ١, ١١) مبرهنة**

لتكن $X_1$، $X_2$، ...و $X_n$ متغيِّرات عشوائيَّة مستمرّة بدوال كثافة احتمالية $f_{X_1}$، $f_{X_2}$، ...و $f_{X_n}$ على الترتيب، فعندئذ إذا كانت دالَّة الكثافة المشتركة $f_{X_1,X_2,\ldots,X_n}$ مُعطاة (أو أمكن تعيينها)، فإنَّ تحقُّق العلاقة الآتية من أجل كلّ $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$:

$$f_{X_1,X_2,\ldots,X_n}(x_1, x_2, \ldots, x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \cdot \ldots \cdot f_{X_n}(x_n)$$

هو **شرط لازم وكاف** لاستقلال المتغيِّرات العشوائيَّة $X_1$، $X_2$، ...و $X_n$.

**(٧, ١, ١, ١٢) مبرهنة**

ليكن $\mathbb{X}_n = (X_1, X_2, \dots, X_n)$ متّجهاً عشوائياً مستمرًّا بدالّة كثافة $f_{\mathbb{X}_n}$، فعندئذ تحقُّق العلاقة الآتية من أجل كلّ $\mathbb{R}^n \ni \vec{x}^n = (x_1, x_2, \dots, x_n)$:

$$f_{\mathbb{X}_n}(\vec{x}^n) = f_1(x_1) \cdot f_2(x_2) \cdot \dots \cdot f_n(x_n) \qquad [7,6]$$

هو **شرط لازم وكاف** لاستقلال المتغيِّرات العشوائيَّة $X_1$، $X_2$،...و $X_n$ المكوِّنة لمركّبات $\mathbb{X}_n$ ، علماً أنَّ $f_1$، $f_2$،...و $f_n$ هي دوال الكثافة الهامشيّة للمتغيِّرات العشوائيَّة $X_1$، $X_2$،...و $X_n$ على الترتيب.

**(٧, ١, ١, ١٣) أمثلة**

١- بالعودة إلى المثال /١/ من (٤-٢-٥-٥) نجد أنَّ المتغيِّرين العشوائيَّين $X$ و $Y$ الذين يكوِّنان مركّبات المتّجه $\mathbb{X}_2 = (X, Y)$ غير مستقلّين بسبب أنّه لدينا:

$$p_{11} = \frac{7}{21} = 0.333 \neq p_{1\bullet} \cdot p_{\bullet 1} = \frac{11}{21} \cdot \frac{14}{21} = \frac{154}{441} = 0.349$$

وبالطبع هذا كاف للحكم على عدم الاستقلال بين مركّبات المتّجه $\mathbb{X}_2 = (X, Y)$.

٢- بالعودة إلى المثال /٢/ من (٤-٢-٥-٥) نجد أنَّ المتغيِّرين العشوائيَّين $X$ و $Y$ الذين يكوِّنان مركّبات المتّجه $\mathbb{X}_2 = (X, Y)$ مستقلّان، وذلك لأنّه من أجل أي $x$ , $y \in \mathbb{R}$ لدينا:

$$f_{\mathbb{X}_2}(x, y) = f_1(x) \cdot f_2(y)$$

علماً أنَّ $f_1$ و $f_2$ هما دالّتي الكثافة الاحتمالية الهامشيتين لـ $X$ و $Y$ على الترتيب. كذلك يمكن استنتاج الاستقلال بين $X$ و $Y$ من خلال تحقُّق العلاقة الآتية أيضاً:

$$F_{\mathbb{X}_2}(x, y) = F_X(x) \cdot F_Y(y) \qquad ; \forall x, y \in \mathbb{R}$$

علماً أنَّ $F_X$ و $F_Y$ هما دالتي التوزيع الهامشيتين لـ $X$ و $Y$ على الترتيب.

**(٧, ١, ٢) الاستقلال لعدد غير منته من المتغيّرات العشوائية**

في هذه الفقرة سنقدِّم مفهوم الاستقلال لعددغير منته من المتغيّرات العشوائية (**ولكن وبشكل موجز**).

**(٧, ١, ٢, ١) تعريف (الاستقلال لعددغير منته من المتغيّرات العشوائية)**

لتكن $T \subseteq \mathbb{R}$ مجموعة غير خالية وليست منتهية (**قد تكون غير قابلة للعدّ**) أيضاً، ولنأخذ $\{X_t\}_{t \in T}$ أسرة متغيّرات عشوائية فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ ، فعندئذ يُقال عن أسرة المتغيّرات العشوائية $\{X_t\}_{t \in T}$ إنَّها **مستقلّة** (**أو مستقلّة في مجموعة الأساس لها** $\Omega$) إذا كان من أجل أيّة مجموعـة جزئية منتهية $\{t_1, t_2, \dots, t_n\} \subseteq T$ لدينا العلاقة الآتية مُحقّقة من أجل أيّة حوادث $A_{t_1}$، $A_{t_2}$، ...و $A_{t_n}$ من $\mathfrak{R}$:

$$P\left(X_{t_1} \in A_{t_1}, X_{t_2} \in A_{t_2}, \dots, X_{t_n} \in A_{t_n}\right) = \prod_{k=1}^{n} P\left(X_{t_k} \in A_{t_k}\right)$$

**(٢ , ٢ , ١ , ٧) مثال**

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، علماً أنَّ:

$$\Omega=[\,0\,,1\,] \quad \& \quad \mathscr{A}=[\,0\,,1\,]\cap\Re \quad \& \quad P(A)=\int_A dx \quad ;\forall A\in\mathscr{A}$$

أي إنَّه لدينا الفضاء الاحتمالي للتوزيع المنتظم المستمر على الفترة $[\,0\,,1\,]$، ولنفترض أنَّ $X$ متغيِّر عشـوائي فوق $[\Omega,\mathscr{A},P]$ يُمثِّل عدداً عشوائياً في نظام العدِّ الثنائي، أي إنَّ $X=0.\alpha_1\,\alpha_2\,\alpha_3\ldots\alpha_n\ldots$ مع $\alpha_k(\omega)\in\{0\,,1\}$ من أجل كلّ $k\in\mathbb{N}$ و $\omega\in\Omega$، ولنثبت استقلال المتغيِّرات العشـوائية $\alpha_1$، $\alpha_2$ و...

من أجل ذلك لنأخذ $n$ متغيِّراً عشوائياً من الأسرة $\{\alpha_k\ ;\ k\in\mathbb{N}\}$ علماً أنَّ $n\in\mathbb{N}$ كيفي ولكنَّه مُثَبَّت، ولنقم بإعادة ترقيم هذه المتغيِّرات العشوائية بالشكل $\alpha_1$، $\alpha_2$ و... $\alpha_n$، ولنثبت استقلال هذه المتغيِّرات العشوائية على النحو الآتي:

من المعلوم أنَّ أي عدد $x\in[0\,,1]$ يُكتب في نظام العدِّ الثنائي وفقاً للعلاقة $x=\sum_{k=1}^{\infty}\alpha_k\,2^{-k}$، ومن ثَمَّ كلّ قيمة $x$ تُعيِّن متتالية متغيِّرات عشوائية $(\alpha_k)_{k\in\mathbb{N}}$، ومنه يكون من أجل أيّة قيم $x_1$، $x_2$ و... $x_n$ من المجموعة $\{0,1\}$ لدينا العلاقة الآتية مُحقَّقة:

$$P\left(\alpha_1=x_1,\alpha_2=x_2,\ldots,\alpha_n=x_n\right)=P\left(x=\sum_{k=1}^{n}x_k\,2^{-k}+\sum_{k=n+1}^{\infty}x_k^{*}\,2^{-k}\quad;\ x_k^{*}\ \ is\ \ arbaitery\right)$$

$$=P\left(\underbrace{\sum_{k=1}^{n}x_k\ 2^{-k}}_{for\ x_i^*=0}\le x\le\underbrace{\sum_{k=1}^{n}x_k\,2^{-k}+\sum_{k=n+1}^{\infty}2^{-k}}_{for\ x_i^*=1}\right)$$

ولكن يمكننا أن نكتب الآتي:

$$\sum_{k=n+1}^{\infty}2^{-k}=\sum_{k=n+1}^{\infty}2^{-(k-n-1)-n-1}=2^{-n-1}\sum_{k=0}^{\infty}2^{-k}=2^{-n-1}\cdot 2=2^{-n}$$

ولذلك ينتج لدينا أنَّ:

$$P\left(\sum_{k=1}^{n}x_k\,2^{-k}\le x\le\sum_{k=1}^{n}x_k\,2^{-k}+\sum_{k=n+1}^{\infty}2^{-k}\right)=\sum_{k=n+1}^{\infty}2^{-k}=2^{-n}$$

وهكذا أصبح لدينا:

$$P\left(\alpha_1=x_1\,,\,\alpha_2=x_2\,,\ldots,\,\alpha_n=x_n\right)=2^{-n} \qquad \textbf{(1)}$$

ومن جهة أخرى نجد أنَّه من أجل كلّ $k\in\mathbb{N}$ لدينا العلاقة $P\left(\alpha_k=x_k\right)=\frac{1}{2}$ مُحقَّقة من أجل أيّ $x_k\in\{0,1\}$، ومن ثَمَّ يكون:

$$\prod_{k=1}^{n}P\left(\alpha_k=x_k\right)=\left(\frac{1}{2}\right)^n=2^{-n} \qquad \textbf{(2)}$$

وبالتالي نجد من (1) و (2) أنَّ:

$$P\left(\alpha_1=x_1\,,\,\alpha_2=x_2\,,\ldots,\alpha_n=x_n\right)=\prod_{k=1}^{n}P\left(\alpha_k=x_k\right)$$

وهذا يعني أنَّ المتغيِّرات العشوائية $\alpha_1$، $\alpha_2$ و... $\alpha_n$ مستقلّة من أجل أيّ $n\in\mathbb{N}$، وبالنتيجة تكون أسرة المتغيِّرات العشوائية $\{\alpha_k\ \ ;k\in\mathbb{N}\}$ مستقلّة أيضاً، وبهذا يكون قد تمَّ إثبات الاستقلال.

## (٧,٢) العزوم للمتغيرات العشوائية

**Moments of random variables**

لقد لاحظنا في الفصل الثاني أنَّ أيّة مجموعة من البيانات سيكون لها عدّة سمات عدديّة تميّزها منها مقاييس النزعة المركزية، التشتت والشكل. كذلك الأمر بالنسبة إلى المتغيِّرات العشوائية، فلها سمات عدديّة تميّزها أيضاً، ولكن بسبب أنَّ اهتمامنا في الفصلين السابقين كان منصبًّا على معرفة نوع المتغيِّر العشوائي وتعيين توزيعه الاحتمالي، وحساب احتمالات لحوادث متعلّقة بالمتغيِّر العشوائي دون الاهتمام بسلوك القيم التي يأخذها المتغيِّر العشوائي، فإنَّنا سنقوم بتقديم بعض المقاييس العددية للمتغيِّرات العشوائية التي تُعدّ بمثابة سمات مميّزة للمتغيِّرات العشوائية أيضاً.

من جهة أخرى قد يكون تعيين دالّة التوزيع الاحتمالية لمتغيِّر عشوائي عسيراً أو غير ممكن وفقاً للتعريف المُعطى لدالّة التوزيع الاحتمالية، ولهذا السبب كان لا بدَّ من البحث عن طرائق أخرى تساعدنا في ذلك، وفي هذا الصدد لوحظ أنَّ ما يُعرف باسم الدوال المولِّدة الاحتمالية (التي ينتج عنها كحالة خاصة الدوال المولِّدة للعزوم) وكذلك الدوال المميِّزة أثبتتا فعَّالية كبيرة في مثل هذه الحالات. لكنَّ مفهوم الدوال المولِّدة وكذلك الدوال المميِّزة يعتمد في تعريفه على مفهوم العزوم للمتغيِّر العشوائي أيضاً، ولذلك كان من الضرورة تقديم مفهوم العزوم للمتغيِّرات العشوائية أولاً، والتعرف على أهمّ خصائصها، ومن ثمَّ الانتقال إلى مفهومي الدوال المولِّدة والدوال المميِّزة للمتغيِّرات العشوائية.

قبل الانتقال إلى تقديم العزوم نشير هنا إلى أنَّنا سنقوم بحساب بعض القيم المميِّزة لبعض المتغيِّرات العشوائية الشهيرة التي ذكرناها في هذا الكتاب، ومن يودّ الاطلاع على بقية القيم المميِّزة التي سنتعامل معها من أجل بقية المتغيِّرات العشوائية الشهيرة السابق ذكرها يمكنه الرجوع إلى جدول لبعض التوزيعات الاحتمالية الشهيرة المقدَّم في نهاية هذا الكتاب.

### (٧,٢,١) العزم الابتدائي من المرتبة $k$ لمتغيِّر عشوائي Moment of Order $k$ of a Random Variable

إنَّ أول سمة مميّزة لمتغيِّر عشوائي هي ما يُعرف باسم **القيمة الوسطى** للمتغيِّر العشوائي التي سنشرحها من أجل متغيِّر عشوائي متقطِّع بسيط وذلك على سبيل التبسيط والتوضيح.

#### (٧,٢,١,١) مفهوم القيمة الوسطى لمتغيِّر عشوائي

ليكن $X$ متغيِّراً عشوائياً متقطِّعاً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ بمجموعة قيم $\mathbf{X}=\{x_i\ ; i\in I\}$ مع $I$ مجموعة منتهية (أي إنَّ $X$ متغيِّر عشوائي بسيط)، فعندئذ بفرض أنَّ:

$$Z_i=\{\omega\in\Omega\ ;\ X(\omega)=x_i\}\qquad ;\forall i\in I$$

فإنَّه (وكما بيّنا في الفصل الخامس) يمكننا أن نكتب:

$$X(\omega)=\sum_{i\in I}x_i\cdot\mathbf{I}_{Z_i}(\omega)\qquad ;\forall\ \omega=\Omega$$

فعندئذ لو قمنا بتكرَّار التجربة العشوائية التي قيد الدرس (وتحت الشروط نفسها، وبشكل مستقل كلّ منها عن الأخرى) لـ $n$ مرّة متتالية، فإنَّنا سنتوقَّع ظهور القيمة $x_i$ عدداً من المرّات يساوي على وجه التقريب $n\cdot p_i$ وذلك من أجل كلّ $i\ni I$ (علماً أنَّ $p_i=P(X=x_i)$)، ومن ثم القيمة الوسطى (أو المتوسط، أو القيمة المتوقَّع الحصول عليها) لقيم هذا المتغيِّر العشوائي المبنية على نتائج الـ $n$ تجربة المنفَّذة هو على وجه التقريب:

$$\frac{x_1\,n\,p_1+x_2\,n\,p_2+\ldots+x_{|I|}\,n\,p_{|I|}}{n}=\sum_{i\in I}x_i\ p_i$$

وهنا نلاحظ أنَّ هذه النتيجة جاءت منسجمة مع مفهوم المتوسط الموزون الذي قُدِّم في الفصل الثاني، ولكن هنا من أجل متغيّر عشوائي،

حيث لدينا الأوزان $w_1 = n \cdot p_1$، $w_2 = n \cdot p_2$،...و $w_{|I|} = n \cdot p_{|I|}$ لـ $x_1$، $x_2$،...و $x_{|I|}$ على الترتيب. إذاً، فالقيمة الوسطى لمتغيِّر عشوائي تُعبِّر عن مركز ثقل مجموعة قيمه أيضاً.

في الواقع إنَّ القيمة الوسطى لمتغيِّر عشوائي هي حالة خاصَّة ممّا يُدعى بـ العزوم لمتغيِّر عشوائي، ولذلك سنقوم أولاً بتقديم مفاهيم تتعلَّق بالعزوم لمتغيِّر عشوائي والتعرف على أهمِّ الخصائص المتعلِّقة بها.

### (٧,٢,١,٢) تعريف العزم الابتدائي من المرتبة $k$ لمتغيِّر عشوائي

ليكن $X$ متغيِّراً عشـوائياً فوق فضاء احتمالي $[\Omega, \mathcal{A}, P]$، و $k$ عدداً صحيحاً موجباً، فعندئذ **العزم الابتدائي من المرتبة $k$** للمتغيِّر العشـوائي $X$ (سنرمز له بـ $\mu_X^{(k)}$) يُحسب على النحو الآتي:

١- إذا كان $X$ **متقطِّعاً** بمجموعة قيم $\mathbf{X} = \{x_i \; ; i \in I\}$ فإنَّ:

$$\mu_X^{(k)} = \sum_{i \in I} x_i^k \cdot P(X = x_i) \qquad [7,7\text{-a}]$$

٢- إذا كان $X$ **مستمراً** بدالَّة كثافة احتمالية $f_X$ فإنَّ:

$$\mu_X^{(k)} = \int_{-\infty}^{+\infty} x^k f_X(x)\, dx \qquad [7,7\text{-b}]$$

### (٧,٢,١,٣) ملاحظات

١- في حال كانت الدراسة متعلِّقة بمتغيِّر عشوائي واحد فقط (ومن ثَمَّ لا مجال للالتباس في هذه الحالة)، فإنَّنا سنكتب $\mu^{(k)}$ عوضاً عن $\mu_X^{(k)}$ وذلك على سبيل التبسيط.

٢- عندما يكون $k = 1$ سنكتب $\mu := \mu^{(1)}$، وفي هذه الحالة الخاصّة يُقال عن $\mu$ إنَّه **التوقع الرياضياتي** Mathematical Expectation للمتغيِّر العشـوائي $X$، وعادةً يُرمز له بـ $\mathbf{E}X$ أيضاً (استخدم الرمز **E** كأول حرف من كلمة **Expectation**)، أي إنَّ $\mu = \mathbf{E}X$، وقياساً عليه يُقال عن $\mathbf{E}X^k$ إنَّه التوقع الرياضياتي للمتغيِّر العشـوائي $X^k$، ومنه يكون $\mu^{(k)} = \mathbf{E}X^k$، ومن ثَمَّ يُمكننا أن نكتب $\mu^{(k)}$ على النحو الآتي:

$$\mu^{(k)} = \mathbf{E}X^k = \begin{cases} \sum\limits_{i \in I} x_i^k \cdot P(X = x_i) & \text{for } X \text{ discrete} \\ \int\limits_{\mathbb{R}} x^k \cdot f_X(x)\, dx & \text{for } X \text{ continuous} \end{cases} \qquad [7,7\text{-c}]$$

٣- كما هو ملاحظ فإنَّ $\mathbf{E}X^k$ هو قيمة عددية، وقد تكون قيمته تساوي $\pm\infty$ عندما يكون $\sum\limits_{i \in I} x_i^k P(X = x_i)$ متباعداً أو $\int_{-\infty}^{+\infty} x^k f_X(x)\, dx$ متباعداً، ولذلك يُقال إنَّ المتغيِّر العشـوائي $X$ **يَملك** عزماً ابتدائياً من المرتبة $k$ إذا كان:

$$\mathbf{E}|X|^k < +\infty \qquad [7,8]$$

٤- إذا كانت $g$ دالَّة حقيقية معرَّفة على $\mathbb{R}$ وقيوسةً بالنسبة إلى $\Re$، فعندئذ تكون الدالَّة $g(X)$ متغيِّراً عشـوائيًّا أيضاً، ولذلك يقال عن $\mathbf{E}[g(X)]^k$ إنَّه العزم الابتدائي من المرتبة $k$ لـ $g(X)$ أو التوقع الرياضياتي لـ $g^k(X)$، حيث يكون لدينا:

**a)** $$\mathbf{E}[g(X)]^k = \sum_{i \in I} [g(x_i)]^k \cdot P(X = x_i) \qquad [7,9\text{-a}]$$

إذا كان $X$ متغيِّراً عشوائياً متقطِّعاً بمجموعة قيم $\mathbf{X} = \{x_i \; ; \; i \in I\}$.

$$\textbf{b)}\quad \mathbf{E}\left[g(X)\right]^k = \int_{-\infty}^{+\infty} \left[g(x)\right]^k \cdot f_X(x)\,dx \qquad [7,9\text{-b}]$$

إذا كان $X$ متغيّراً عشوائياً مستمراً بدالّة كثافة احتمالية $f_X$.

أمّا وجودية $\mathbf{E}\left[g(X)\right]^k$ فإنّه مرهونة بالتقارب المطلق للمتسلسلة العددية [7,9-a] أو للتكامل المعتل [7,9-b] ، وهنا نلاحظ أنّه في الحالة الخاصّة عندما تكون $g$ دالّة المطابقة على $\mathbb{R}$ (أي إنّ $g(x) = x$)، فإنّه يصبح لدينا $\mathbf{E}X = \mathbf{E}[g(X)]$ هو التوقُّع الرياضياتي للمتغيِّر العشوائي $X$ نفسه.

٥- يُبرهن (سنقبلها دون برهان) على أنّه إذا كان $X$ متغيِّراً عشوائياً **يملك** عزماً ابتدائياً من المرتبة $k$ مع $k \in \mathbb{N}$ (أي إنّ $\mu^{(k)}$ موجوداً)، فإنّ العزوم الابتدائية $\mu^{(\ell)}$ ستكون موجودةً من أجل كلّ قيم $\ell$ المُحقِّقة للمتباينة $1 \le \ell < k$، وتكون المتباينة الآتية مُحقَّقةً أيضاً:

$$\sqrt[\ell]{\mathbf{E}|X|^{\ell}} \le \sqrt[k]{\mathbf{E}|X|^{k}}$$

وبناءً على هذه المقولة، فإذا كان $X$ متغيِّراً عشوائياً **لا يملك** عزماً ابتدائياً من المرتبة $k$ فإنّه لن يملك أي عزم ابتدائي من مرتبة $k < m$ .

### (٤, ١, ٢, ٧) أمثلة

١- ليكن $[\Omega,\mathcal{A},P]$ الفضاء الاحتمالي الموافق لتجربة إلقاء حجر نرد متوازن لمَرّة واحدة فقط، وليكن $X$ متغيِّراً عشوائياً فوق هذا الفضاء الاحتمالي مُعرَّفاً من خلال العلاقة $X(\omega) = \omega$ ، ولنبيّن إن كان هذا المتغيّر العشوائي يملك عزوماً ابتدائية من المرتبة الأولى والثانية أم لا؟

الحـل: من أجل ذلك نلاحظ أنّ هذا المتغيِّر العشوائي متقطِّع ومجموعة قيمه هي $\mathbf{X} = \{1,2,3,4,5,6\}$ ، ومن ثَمَّ يكون لدينا:

$$\mathbf{E}X = \sum_{i\in I} x_i \cdot P(X = x_i) = \sum_{k=1}^{6} k \cdot P(X = k) = \sum_{k=1}^{6} k\,\frac{1}{6} = \frac{1}{6}\sum_{k=1}^{6} k = \frac{21}{6} = 3.5$$

وكذلك نجد أنَّ:

$$\mathbf{E}X^2 = \sum_{i\in I} x_i^2 \cdot P(X = x_i) = \sum_{k=1}^{6} k^2 \cdot P(X = k) = \frac{1}{6}\sum_{k=1}^{6} k^2 = \frac{91}{6} = 15.17$$

وبحساب العزوم المطلقة نجد ما يلي:

$$\mathbf{E}|X| = \sum_{i\in I} |x_i| \cdot P(X = x_i) = \sum_{k=1}^{6} |k| \cdot P(X = k) = \frac{1}{6}\sum_{k=1}^{6} k = 3.5 < +\infty$$

ومن ثَمَّ المتغيِّر العشوائي $X$ يملك عزماً ابتدائياً من المرتبة الأولى (يملك توقعاً رياضياتياً)، وكذلك لدينا:

$$\mathbf{E}|X|^2 = \sum_{i\in I} |x_i|^2 \cdot P(X = x_i) = \sum_{k=1}^{6} |k|^2 \cdot P(X = k) = \frac{1}{6}\sum_{k=1}^{6} k^2 = 15.17 < +\infty$$

ومن ثَمَّ المتغيِّر العشوائي $X$ يملك عزماً ابتدائياً من المرتبة الثانية أيضاً، وبشكل مماثل يمكننا تبيان أنَّ $X$ يملك عزوماً ابتدائيةً من مختلف المراتب $k \in \mathbb{N}$ .

٢- تحت فرضيات المثال السابق / ١/ سنأخذ $g$ دالّة حقيقية معرَّفة على $\mathbb{R}$ من خلال العلاقة $g(x) := 2x + 1$ ، ولنبيِّن إن كان المتغيّر العشوائي $Y = g(X)$ يملك عزماً ابتدائيًّا من المرتبة الأولى أم لا؟

الحـل: من الملاحظ أنَّ $g(X)=2X+1$ هو متغيِّر عشوائي متقطِّع أيضاً لأنَّ الدالَّة المُعطاة $g$ مستمرَّة على $\mathbb{R}$، ولدينا $X$ متغيِّر عشوائي متقطِّع بالفرض، ولذلك يمكننا أن نكتب الآتي:

$$\mathbf{E}Y=\mathbf{E}\left[g(X)\right]=\sum_{i\in I}g(x_i)P\left(X=x_i\right)=\sum_{k=1}^{6}(2k+1)\frac{1}{6}=\frac{1}{6}\sum_{k=1}^{6}(2k+1)=\frac{48}{6}=8$$

ومن جهة أخرى لدينا:

$$\mathbf{E}\left|Y\right|=\mathbf{E}\left|g(X)\right|=\sum_{i\in I}\left|g(x_i)\right|\cdot P\left(X=x_i\right)=\sum_{k=1}^{6}\left|2k+1\right|\frac{1}{6}=\frac{1}{6}\sum_{k=1}^{6}(2k+1)=\frac{48}{6}=8<+\infty$$

ومن ثَّم المتغيِّر العشوائي $Y=g(X)$ يملك عزماً ابتدائياً من المرتبة الأولى (يملك توقعاً رياضياتياً)، وبشكل مماثل يمكننا تبيان أنَّ $Y=g(X)$ يملك عزوماً ابتدائيةً من مختلف المراتب $k\ni\mathbb{N}$.

٣- ليكن $X$ متغيِّراً عشوائياً بحيث $X\sim\mathbf{U}\left([a,b]\right)$، ولنبيِّن إن كان هذا المتغيِّر العشوائي يملك عزوماً ابتدائيةً من المرتبة الأولى والثانية أم لا؟

الحـل: نعلم أنَّ هذا المتغيِّر العشوائي مستمرّ بدالَّة كثافة احتمالية:

$$f_X(x)=\frac{1}{b-a}\mathbf{I}_{[a,b]}(x)\qquad;\forall x\in\mathbb{R}$$

مع $-\infty<a<b<+\infty$ ثوابت، ومن ثَّم يكون لدينا:

$$\mathbf{E}(X)=\int_{-\infty}^{\infty}x\cdot f_X(x)\,dx=\int_a^b\frac{1}{b-a}x\,dx=\frac{1}{b-a}\left[\frac{b^2-a^2}{2}\right]=\frac{a+b}{2}$$

وكذلك نجد:

$$\mathbf{E}\left|X\right|=\int_{-\infty}^{\infty}\left|x\right|\cdot f_X(x)\,dx=\frac{1}{b-a}\int_a^b\left|x\right|dx<+\infty$$

وذلك لأنَّه:

أ- من أجل $0<a$ لدينا:

$$\int_a^b\left|x\right|dx=\int_a^b x\,dx=\frac{b^2-a^2}{2}<+\infty$$

ب- من أجل $0>a$ و $0<b$ لدينا:

$$\int_a^b\left|x\right|dx=\int_a^0 -x\,dx+\int_0^b x\,dx=\frac{a^2+b^2}{2}<+\infty$$

ج- من أجل $0>b$ لدينا:

$$\int_a^b\left|x\right|dx=\int_a^b -x\,dx=\frac{a^2-b^2}{2}<+\infty$$

ومن ثَّم المتغيِّر العشوائي $X$ يملك عزماً ابتدائياً من المرتبة الأولى (يملك توقعاً رياضياتياً)، وكذلك نجد أنَّ المتغيِّر العشوائي $X$ يملك عزماً ابتدائياً من المرتبة الثانية أيضاً، حيث لدينا:

$$\mathbf{E}\left|X\right|^2=\int_{-\infty}^{\infty}\left|x\right|^2\cdot f_X(x)\,dx=\frac{1}{b-a}\int_a^b x^2\,dx=\frac{b^3-a^3}{3(b-a)}=\frac{(b^2+ab+a^2)}{3}<+\infty$$

وبشكل مماثل يمكننا تبيان أنَّ $X$ يملك عزوماً ابتدائيةً من مختلف المراتب $k\ni\mathbb{N}$.

٤- تحت فرضيات المثال السابق /٣/ سنأخذ $g$ دالّة حقيقية معرَّفة على $\mathbb{R}$ من خلال العلاقة $g(x):=2x+1$، ولنبيّن إن كان المتغيّر العشوائي $Y=g(X)$ يملك عزماً ابتدائيًا من المرتبة الأولى أم لا؟

الحـل: نلاحظ أنَّ المتغيِّر العشوائي $Y$ مستمرّ لأنَّ الدالَّة $g(x):=2x+1$ مستمرّة ،ولدينا $X$ متغيِّر عشوائي مستمرّ بالفرض، ومن ثمَّ يكون:

$$\mathbf{E}[Y]=\mathbf{E}[g(X)]=\int_{-\infty}^{\infty} g(x)\cdot f_X(x)\,dx=\int_a^b \frac{1}{b-a}(2x+1)\,dx=\frac{1}{b-a}\Big[\left(b^2-a^2\right)+\left(b-a\right)\Big]=b+a+1$$

وبما أنَّ $-\infty<a<b<+\infty$ فإنَّه سينتج لدينا أنَّ $\mathbf{E}\,|g(X)|<+\infty$، وهذا يعني أنَّ المتغيِّر العشوائي $Y=2X+1$ يملك عزماً ابتدائياً من المرتبة الأولى (يملك توقعاً رياضياً)، وبشكل مماثل يمكننا تبيان أنَّ $Y$ يملك عزوماً ابتدائيةً من مختلف المراتب $k\in\mathbb{N}$.

أما إذا أردنا حساب $\mathbf{E}(Y)$ من خلال تعيين $f_Y(y)$ فإنَّنا نجد من أجل كلّ $y$ من $\mathbb{R}$ ما يلي:

$$F_Y(y)=F_X\left(\frac{y-1}{2}\right)=\int_{-\infty}^{\frac{y-1}{2}}\frac{1}{b-a}\,\mathbf{I}_{[a,b]}(t)\,dt=\int_{-\infty}^{y}\underbrace{\frac{1}{2(b-a)}\,\mathbf{I}_{[a,b]}\left(\frac{y-1}{2}\right)}_{f_Y(y)}dy$$

والتي ينتج عنها أنَّ:

$$f_Y(y)=\frac{1}{2(b-a)}\,\mathbf{I}_{[a,b]}\left(\frac{y-1}{2}\right)$$

ومن ثم يكون لدينا:

$$\mathbf{E}\,Y=\int_{-\infty}^{\infty} y\cdot f_Y(y)\,dy=\frac{1}{2(b-a)}\int_{-\infty}^{\infty} y\cdot\mathbf{I}_{[a,b]}\left(\frac{y-1}{2}\right)dy=\frac{1}{2(b-a)}\int_{-\infty}^{\infty} y\cdot\mathbf{I}_{[2a+1,\,2b+1]}(y)\,dy$$

$$=\frac{1}{2(b-a)}\left[\underbrace{\int_{-\infty}^{2a+1} y\cdot\overbrace{\mathbf{I}_{[2a+1,\,2b+1]}(y)}^{=0}\,dy}_{=0}+\int_{2a+1}^{2b+1} y\cdot\overbrace{\mathbf{I}_{[2a+1,\,2b+1]}(y)}^{=1}\,dy+\underbrace{\int_{2b+1}^{+\infty} y\cdot\overbrace{\mathbf{I}_{[2a+1,\,2b+1]}(y)}^{=0}\,dy}_{=0}\right]$$

$$=\frac{1}{2(b-a)}\int_{2a+1}^{2b+1} y\,dy=\frac{1}{2(b-a)}\cdot\frac{y^2}{2}\Bigg|_{2a+1}^{2b+1}=\frac{1}{2(b-a)}\cdot\frac{4(b^2-a^2)+4(b-a)}{2}$$

ومنه ينتج لدينا الآتي (وهي ذات النتيجة التي حصلنا عليها سابقاً):

$$\mathbf{E}\,Y=\frac{1}{2(b-a)}\cdot\frac{4(b-a)\,[(b+a)+1]}{2}=b+a+1$$

٥- ليكن $X\sim\mathbf{C}(0,1)$ (خاضع توزيع كوشي)، ولنبيّن إن كان هذا المتغيّر العشوائي يملك عزماً ابتدائياً من المرتبة الأولى أم لا؟

الحـل: نعلم أنَّ لهذا المتغيّر العشوائي دالَّة كثافة احتمالية تُعطى من خلال العلاقة الآتية (انظر العلاقة [6-14-a]):

$$f_X(x)=\frac{1}{\pi(1+x^2)}\qquad;\forall x\in\mathbb{R}$$

ومن ثمَّ يكون لدينا:

$$\mathbf{E}\,|X| = \int_{-\infty}^{\infty} |x|\, f_X(x)dx = \int_{-\infty}^{\infty} |x| \frac{1}{\pi\,(1+x^2)}\; dx = \frac{1}{\pi}\left[\int_{-\infty}^{0} \frac{-x}{1+x^2}\; dx + \int_{0}^{\infty} \frac{x}{1+x^2}\, dx\right]$$

$$= \frac{1}{\pi}\int_{0}^{\infty} \frac{2x}{1+x^2}\, dx \overset{y=1+x^2}{=\!=\!=\!=\!=} \frac{1}{\pi}\int_{1}^{\infty} \frac{dy}{y} = \frac{1}{\pi}\left[\ln y \,\Big|_{1}^{\infty}\right] = +\infty$$

أي إنَّ $\mathbf{E}\,|X| = +\infty$، وهذا يعني أنَّ المتغيِّر العشوائي الخاضع لتوزيع كوشي لا يملك عزماً ابتدائياً من المرتبة الأولى، ومن ثمَّ فإنَّه **لا يملك** عزوماً ابتدائيةً ولا من أيَّة مرتبة.

**(٧, ٢, ١, ٥) ملاحظة**

من المثال الأول في الفقرة السابقة وجدنا $\mathbf{E}X = \frac{21}{6}$، ومن ثَمَّ يكون لدينا:

$$g\left(\mathbf{E}X\right) = g\left(\frac{21}{6}\right) = 2\frac{21}{6} + 1 = \frac{48}{6} = 8$$

ومن ثَمَّ يكون لدينا من أجل هذا المثال ما يلي مُحقَّقاً:

$$\mathbf{E}\left[g\left(X\right)\right] = g\left(\mathbf{E}\,X\right)$$

ولكن يجب الانتباه إلى أنَّه في الحالة العامَّة لدينا:

$$\mathbf{E}\left[g\left(X\right)\right] \neq g\left(\mathbf{E}\,X\right)$$

ولكن عندما تتمتَّع الدالَّة $g$ ببعض الخصائص فثمَّة علاقة مُحدَّدة بين طرفي العلاقة السابقة، وهذا ما تقدِّمه لنا المبرهنة الآتية التي سنقبلها دون برهان.

**(٧, ٢, ١, ٦) مبرهنة (متباينة جينسين Jensen inequality)**

لتكن $g$ دالَّة حقيقية مستمرَّة على $\mathbb{R}$، وليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $\left[\Omega, \mathscr{A}, P\right]$، فعندئذ:

١- إذا كانت الدالَّة $g$ محدَّبة على $\mathbb{R}$ فإنَّه يكون لدينا:

$$\mathbf{E}\left[g\left(X\right)\right] \geq g\left(\mathbf{E}X\right) \qquad [7,10\text{-a}]$$

٢- إذا كانت الدالَّة $g$ مقعَّرة على $\mathbb{R}$ فإنَّه سيكون لدينا:

$$\mathbf{E}\left[g\left(X\right)\right] \leq g\left(\mathbf{E}X\right) \qquad [7,10\text{-b}]$$

إنَّ كلاً من المتباينتين السابقتين تُعرف باسم متباينة جينسين نسبة إلى الرياضياتي والمهندس الدانماركي جينسين Johan Ludwig William Valdemar Jensen (1859-1925).

لاحظ أنَّ المساواة تتحقَّق من أجل بعض الحالات في كلا المتباينتين السابقتين، ومنها الحالة التي تكون فيها $g$ مُعطاة من خلال العلاقة $y = g\left(x\right) = 2x + 1$ التي عرضناها سابقاً حيث نعلم أنَّ هذه الدالَّة تُحقِّق تعريف الدالَّة المحدَّبة والمقعَّرة بآن واحد (من أجل الاطلاع على تعريف الدالَّة المحدَّبة والمقعَّرة راجع الملحق A).

**(٧, ٢, ١, ٧) أمثلة**

١- لنأخذ $X$ متغيِّراً عشوائياً كما في المثال /١/ من الفقرة (٧, ٢, ١, ٤)، ولنأخذ $g$ دالَّة حقيقية معرَّفة على $\mathbb{R}$ من خلال العلاقة $g\left(x\right) = \mathbf{e}^{2x}$، فنلاحظ أنَّ الدالَّة $g$ مستمرة على $\mathbb{R}$، ولذلك فإنَّ $g\left(X\right)$ هو متغيِّر عشوائي (متقطِّع لأنَّ $X$ متقطِّع)، وكذلك نجد أنَّ $g$ محدَّبة على $\mathbb{R}$، ومن ثَمَّ يمكننا أن نكتب الآتي:

$$\mathbf{E}\left[g(X)\right]=\sum_{i\in I}g(x_i)P\left(X=x_i\right)=\sum_{k=1}^{6}\mathbf{e}^{2k}\ \frac{1}{6}\ =\frac{1}{6}\sum_{k=1}^{6}\mathbf{e}^{2k}=\frac{188227.631}{6}=31371.272$$

ولكن وجدنا سابقاً $\mathbf{E}\,X = 3.5$ ، ومن ثمَّ يكون لدينا $g\left[\mathbf{E}(X)\right]=\mathbf{e}^7=1096.333$ ، والتي نجد منها أنَّ:

$$g\ \left[\mathbf{E}(X)\right]<\mathbf{E}\left[g(X)\right]$$

٢- مرَّةً أخرى لنأخذ $X$ متغيِّراً عشوائياً كما في المثال / ١/ من الفقرة (٧ , ٢ , ١ , ٤)، وسنأخذ $g$ دالَّة حقيقية معرَّفة على $\mathbb{R}^+$ من خلال العلاقة $g(x)=\sqrt{x}$ ، فنجد أنَّ الدالَّة $g$ مستمرَّة ومقعَّرة على $\mathbb{R}^+$ ، فينتج بسبب استمرار الدالَّة $g$ أنَّ $g(X):=\sqrt{X}$ متغيِّر عشوائي متقطِّع، ومن ثمَّ يمكننا أن نكتب الآتي:

$$\mathbf{E}\left[g(X)\right]=\sum_{i\in I}g(x_i)P\left(X=x_i\right)=\sum_{k=1}^{6}\sqrt{k}\ \frac{1}{6}\ =\frac{1}{6}\sum_{k=1}^{6}\sqrt{k}\ =\frac{10.832}{6}=1.805$$

ولكن لدينا $\mathbf{E}\,X = 3.5$ ، ومن ثمَّ يكون لدينا $g\ \left[\mathbf{E}(X)\right]=\sqrt{3.5}=1.871$ ، وبالتالي يُختم المثال بالنتيجة الآتية:

$$g\ \left[\mathbf{E}(X)\right]>\mathbf{E}\left[\ g(X)\right]$$

من العلاقات الشهيرة بين العزوم الابتدائية والعزوم الارتباطية لمتغيِّرين عشوائيين ما يُعرف باسم متباينة كوشي- شفارتز التي تقدِّمها المبرهنة الآتية (سنقبلها دون برهان).

**(٧ , ٢ , ١ , ٨) مبرهنة (متباينة كوشي- شفارتز Cauchy – Schwartz inequality)**

ليكن $X$ و$Y$ متغيِّرين عشوائيين فوق فضاء احتمالي $\left[\Omega,\mathcal{A},P\right]$ بحيث إنَّ $\mathbf{E}X^2<+\infty$ وكذلك $\mathbf{E}Y^2<+\infty$ ، فعندئذ تكون المتباينة الآتية مُحقَّقة من أجلهما:

$$\left[\mathbf{E}(XY)\right]^2\leq\mathbf{E}X^2\cdot\mathbf{E}Y^2 \qquad [7,11]$$

وهذه المتباينة تعرف باسم متباينة كوشي- شفارتز (كارل هيرمان أمادوس شفارتز فيلسوف ورياضياتي ألماني Karl Hermann Amandus Schwarz (1843-1921))، وتعدّ هذه العلاقة من العلاقات المهمّة في مجالات عديدة من الرياضيات (الاحتمالات، التحليل، الجبر و...).

**(٧ , ٢ , ١ , ٩) مثال:**

ليكن $\left[\Omega,\mathcal{A},P\right]$ الفضاء الاحتمالي الموافق لتجربة إلقاء حجر نرد متوازن لمرَّة واحدة فقط، ولنفترض أنَّ $X$ و$Y$ متغيِّران عشوائيان مستقلَّان فوق $\left[\Omega,\mathcal{A},P\right]$، ومُعرَّفان من خلال العلاقتين $X(\omega)=\omega$ و $Y(\omega)=\omega^2$، فنجد أنَّ مجموعة قيم كلّ من المتغيِّرين العشوائيين $X$ و$Y$ على الترتيب هما $\mathbf{X}=\{1,2,3,4,5,6\}$ و $\mathbf{Y}=\{1,4,9,16,25,36\}$ ، ولكن وجدنا سابقاً (في المثال / ١/ من (٧ , ٢ , ١ , ٤)) أنَّ $\mathbf{E}X=3.5$ و $\mathbf{E}X^2=15.1667$ ، وكذلك لدينا:

$$\mathbf{E}\,Y=\sum_{j\in J}y_j\cdot P(Y=y_j)=\sum_{\ell=1}^{6}\ell\cdot P\left(Y=\ell\right)=\sum_{\ell=1}^{6}\ell\frac{1}{6}=\frac{1}{6}\sum_{\ell=1}^{6}\ell=\frac{91}{6}=15.167$$

ومن أجل العزم الابتدائي من المرتبة الثانية نجد:

$$\mathbf{E}\ Y^2=\sum_{j\in J}y_j^2\cdot P\left(Y=y_j\right)=\sum_{\ell=1}^{6}\ell^2\cdot P\left(X=\ell\right)=\frac{1}{6}\sum_{\ell=1}^{6}\ell^2=\frac{2275}{6}=379\ .167$$

ومن ثمَّ يكون لدينا $\mathbf{E}\,X^2\cdot\mathbf{E}\,Y^2=5750.826$ ، ومن جهة أخرى يمكننا أن نكتب بسبب الاستقلال بين $X$ و$Y$ ما يلي:

$$\mathbf{E}(XY) = \sum_{i\in I, j\in J}(x_i \cdot y_j)\cdot P\left(X = x_i, Y = y_j\right) = \sum_{i\in I, j\in J} x_i \cdot y_j \cdot P\left(X = x_i\right)\cdot P\left(Y = y_j\right)$$
$$= \sum_{i\in I} x_i \cdot P\left(X = x_i\right) \sum_{j\in J} y_j P\left(Y = y_j\right) = \mathbf{E}(X)\cdot\mathbf{E}(Y) = 53.0845$$

ومن ثَمّ يكون لدينا:

$$2817.964 = \left[\mathbf{E}(X\cdot Y)\right]^2 < \mathbf{E}X^2\cdot\mathbf{E}Y^2 = 5750.826$$

وهذه النتيجة متوافقة مع متباينة كوشي – شفارتز.

الآن، وقبل تقديم بعض الخصائص للعزوم الابتدائية سنقوم بتقديم مفهوم مهمّ يندرج فيه متغيِّرين عشوائيين بآنٍ واحد.

**(١٠ , ١ , ٢ , ٧) تعريف (التغاير – أو تمام التباين– لمتغيِّرين عشوائيين)**

ليكن $X$ و$Y$ متغيِّرين عشوائيين فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ بحيث إنَّ $\mathbf{E}|X| < +\infty$ و $\mathbf{E}|Y| < +\infty$، فعندئذ تدعى القيمة العددية $\mathbf{E}\left[(X-\mathbf{E}X)\cdot(Y-\mathbf{E}Y)\right]$ بـ **التغاير** (أو تمام تباين) Covariance لـ $X$ و$Y$، ويُرمز له بـ $\mathbf{cov}(X,Y)$. أي إنَّ:

$$\mathbf{cov}(X,Y) = \mathbf{E}\left[(X-\mathbf{E}X)\cdot(Y-\mathbf{E}Y)\right] \qquad [7,12]$$

**(١١ , ١ , ٢ , ٧) أمثلة**

١– ليكن $[\Omega,\mathscr{A},P]$ الفضاء الاحتمالي الموافق لتجربة إلقاء حجر نرد متوازن لمرَّة واحدة فقط، وليكن $X$ و$Y$ متغيِّرين عشوائيين فوق $[\Omega,\mathscr{A},P]$ ومُعرَّفين من خلال العلاقتين $X(\omega) = \omega$ و $Y(\omega) = 3\omega$، فنجد أنَّ مجموعة القيم لكلٍّ من المتغيِّرين العشوائيين $X$ و$Y$ على الترتيب هما $\mathbf{X} = \{1,2,3,4,5,6\}$ و $\mathbf{Y} = \{3,6,9,12,15,18\}$، ولنفترض أنَّ قانون التوزيع المشترك لـ $X$ و$Y$ مُعطى كما في الجدول الآتي:

**الجدول (١ , ٧)**

| $i$ | 1 | 2 | 3 | 4 | 5 | 6 | |
|---|---|---|---|---|---|---|---|
| قيم $X$ $x_i$ | 1 | 2 | 3 | 4 | 5 | 6 | sum |
| قيم $Y$ $y_i$ | 3 | 6 | 9 | 12 | 15 | 18 | |
| $P(X = x_i, Y = y_i)$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | 1 |

وبناء على هذه المُعطيات يجب أن يكون:

$$P\left(X = x_i, Y = y_j\right) = 0 \quad ; i, j \in \mathbb{N}_6 \ with \ i \neq j$$

وبحساب $\mathbf{E}X$ نجد أنَّ $\mathbf{E}X = 3.5$، وكذلك:

$$\mathbf{E}Y = \sum_{j\in J} y_j \cdot P(Y = y_j) = \sum_{\ell=1}^{6} \ell \cdot P\left(Y = \ell\right) = \sum_{\ell=1}^{6} \ell\frac{1}{6} = \frac{1}{6}\sum_{\ell=1}^{6} \ell = \frac{63}{6} = 10.5$$

ومن ثَمّ يكون لدينا تغاير $X$ و$Y$ يساوي:

$$\mathbf{cov}(X,Y) = \mathbf{E}\left[(X-\mathbf{E}X)(Y-\mathbf{E}Y)\right] = \sum_{i\in I, j\in J}(x_i - 3.5)(y_j - 10.5)P\left(X = x_i, Y = y_j\right)$$
$$= \frac{1}{6}\sum_{k=1}^{6}(k-3.5)(3k-10.5) = \frac{(18.75)+(6.75)+(0.75)+(0.75)+(6.75)+(18.75)}{6} = 8.75$$

وبهذا ينتهي حل هذا المثال.

٢- بالعودة إلى المثال (٩ , ١ , ٢ , ٧) نجد ما يلي:

$$\mathbf{E}\left[(X-\mathbf{E}X)(Y-\mathbf{E}Y)\right]=\sum_{i\in I,j\in J}\left[(x_i-3.5)(y_j-15.167)\right]\cdot P\left(X=x_i,Y=y_j\right)$$

والآن، وبسبب الاستقلال بين $X$ و$Y$ يكون كلّ من $X-\mathbf{E}X$ و $Y-\mathbf{E}Y$ مستقلّين أيضاً، ولذلك يمكننا أن نكتب الآتي:

$$\mathbf{cov}(X,Y)=\mathbf{E}\left[(X-\mathbf{E}X)(Y-\mathbf{E}Y)\right]=\sum_{i\in I,j\in J}\left[(x_i-3.5)(y_j-15.167)\right]\cdot P\left(X=x_i\right)P\left(Y=y_j\right)$$
$$=\sum_{i\in I}(x_i-3.5)\cdot P\left(X=x_i\right)\sum_{j\in J}(y_j-15.167)\cdot P\left(Y=y_j\right)=0$$

وهكذا ينتج لدينا:

$$\mathbf{cov}(X,Y)=\mathbf{E}\left[(X-\mathbf{E}X)(Y-\mathbf{E}Y)\right]=0$$

وهنا يجب الانتباه إلى أنَّ $\mathbf{E}(Y-\mathbf{E}Y)=0$ وأيّة قيمة مختلفة عن الصفر (قد تظهر في حساب $\mathbf{E}(Y-\mathbf{E}Y)$) تكون ناتجة عن عمليّة تدوير الأرقام في حساب $\mathbf{E}Y$.

نقدِّم فيما يلي بعض خصائص التوقّع الرياضياتي لمتغيِّر عشوائي، علماً أنَّه إذا أردنا إثبات صحة بعض هذه الخصائص فإنَّنا سنفترض أنَّ $X$ و$Y$ متغيِّرين عشوائيين متقطِّعين وذلك على سبيل التوضيح والتبسيط، وأمّا من أجل $X$ و$Y$ متغيِّرين عشوائيين مستمرين فإنَّ هذه الخصائص تبرهن بطريقة مماثلة ولكن باستخدام التكاملات عوضاً عن المجاميع.

**(١٢ , ١ , ٢ , ٧) مبرهنة**

ليكن $X$ و $Y$ متغيِّرين عشوائيين فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$، وكلّ منهما يملك عزماً ابتدائياً من المرتبة الأولى (التوقُّع الرياضياتي لـ $X$ و$Y$ موجود)، فعندئذ إذا كان من أجل $-\infty<a\le b<+\infty$ لدينا $a\le X\le b$، فإنَّه سيكون:

$$a\le\mathbf{E}X\le b \qquad [7,13]$$

**البرهان:** بما أنَّ $a\le X\le b$ و $X$ يملك توقُّعاً رياضياتياً فإنَّه يمكننا أن نكتب الآتي:

$$a=a\underbrace{\sum_{i\in I}P(X=x_i)}_{=1}=\sum_{i\in I}a\cdot P(X=x_i)\le\underbrace{\sum_{i\in I;a\le x_i\le b}x_i\cdot P(X=x_i)}_{\mathbf{E}X}$$

ولكن لدينا:

$$\mathbf{E}X=\sum_{i\in I;a\le x_i\le b}x_i\cdot P(X=x_i)\le\sum_{i\in I}b\cdot P(X=x_i)=b\underbrace{\sum_{i\in I}P(X=x_i)}_{=1}=b$$

أي إنَّ $a\le\mathbf{E}X\le b$، وبهذا يتمّ الإثبات.

**(١٣ , ١ , ٢ , ٧) نتائج**

من المبرهنة السابقة نستنتج ما يلي:

١- من أجل $a=b$ يصبح لدينا $\mathbf{E}X=\mathbf{E}b=b$.

٢- من أجل $a=0$ يصبح $0\le X$، ونتيجة للخاصّية السابقة ينتج لدينا أنَّه إذا كان $X$ متغيِّراً عشوائياً غير سالب فإنَّه سيكون لدينا $0\le\mathbf{E}X$.

**(٧,٢,١,١٤) مبرهنة**

ليكن $X$ و$Y$ متغيِّرين عشوائيين فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، وكلّ منهما يملك عزماً ابتدائياً من المرتبة الأولى، فإذا كان $a\ ,\ b \in \mathbb{R}$ ثوابت مع $a \neq 0$، فإنَّ العلاقة الآتية ستكون مُحقَّقة:

$$\mathbf{E}\left(aX+bY\right)=a\cdot\mathbf{E}X+b\cdot\mathbf{E}Y \qquad [6,14\text{-a}]$$

**البرهان:** بما أنَّ كلّ من $X$ و$Y$ يملك توقُّعاً رياضياتياً، فعندئذ يمكننا أنَّ نكتب الآتي:

$$\begin{aligned}\mathbf{E}(a\cdot X+b\cdot Y)&=\sum_{i\in I,\,j\in J}(a\cdot x_i+b\cdot y_j)\,P(X+Y=x_i+y_j)\\&=\sum_{i\in I,\,j\in J}[a\cdot x_i\cdot P(X+Y=x_i+y_j\mid y=y_j)+b\cdot y_j\cdot P(X+Y=x_i+y_j\mid X=x_i)]\\&=a\sum_{i\in I}x_iP(X+y_j=x_i+y_j)+b\sum_{j\in J}y_j\cdot P(x_i+Y=x_i+y_j)\\&=a\sum_{i\in I}x_i\ P(X+x_i)+b\sum_{j\in J}y_j\ P(y=y_j)\ =a\,\mathbf{E}X+b\,\mathbf{E}Y\end{aligned}$$

علماً أنَّ الاحتمال الشرطي لـ $X=x_i$ علماً أنَّ $Y=y_j$ مع $0<P(Y=y_j)$ مُحقَّقاً يجب أنَّ يفهم على النحو الآتي:

$$P(X=x_i|Y=y_j)=\frac{P\left(\{\omega\in\Omega\,;X(\omega)=x_i\}\bigcap\{\omega\in\Omega\,;Y(\omega)=y_j\}\right)}{P\left(\{\omega\in\Omega\ ;Y(\omega)=y_j\}\right)}$$

وبهذا يتمّ البرهان.

**(٧,٢,١,١٥) نتائج**

من هذه المبرهنة يتضَّح لنا أنَّ التوقع الرياضياتي هو مؤثِّر خطي، وعلاوةً على ذلك ينتج:

١- من أجل الحالة الخاصّة $P(Y=1)=1$ سيكون لدينا:

$$\mathbf{E}\left(aX+b\right)=a\cdot\mathbf{E}X+b \qquad [7,14\text{-b}]$$

٢- من أجل الحالة الخاصّة $a=1$ و $\mathbf{E}X=b$ سيكون لدينا:

$$\mathbf{E}\left(X-\mathbf{E}X\right)=0 \qquad [7,14\text{-c}]$$

علماً أنَّ المتغيِّر العشوائي $X-\mathbf{E}X$ يُدعى **متغيِّراً عشوائيّاً مركزيّاً**، ومن ثَمَّ يكون التوقُّع الرياضياتي لمتغيِّر عشوائي مركزي يساوي الصفر دوماً (وقد لاحظنا هذا في أمثلة سابقة).

**(٧,٢,١,١٦) مبرهنة**

ليكن $X$ و$Y$ متغيِّرين عشوائيين فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، وكلّ منهما يملك عزماً ابتدائياً من المرتبة الأولى، فعندئذ يكون لدينا:

$$\mathbf{E}\left(X\cdot Y\right)=\mathbf{E}X\cdot\mathbf{E}Y-\mathbf{cov}(X,Y) \qquad [7,15\text{-a}]$$

**البرهان:** بما أنَّ كلًّا من $X$ و$Y$ يملك توقُّعاً رياضياتياً، فعندئذ بملاحظة أنَّ:

$$XY=\mathbf{E}X\cdot\mathbf{E}Y+(X-\mathbf{E}X)\,\mathbf{E}Y+(Y-\mathbf{E}Y)\,\mathbf{E}X+(X-\mathbf{E}X)\cdot(Y-\mathbf{E}Y)$$

وباستخدام المبرهنتين السابقتين ونتائجهما تنتج لدينا العلاقة الآتية التي تثبت صحة العلاقة [7,15-a]:

$$\mathbf{E}(XY)=\mathbf{E}X\cdot\mathbf{E}Y+\mathbf{E}\left[(X-\mathbf{E}X)\,(Y-\mathbf{E}Y)\right]=\mathbf{E}X\cdot\mathbf{E}Y+\mathbf{cov}(XY)$$

**(٧, ٢, ١, ١٧) مبرهنة**

ليكن $X$ و$Y$ متغيِّرين عشوائيين فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$، وكلّ منهما يملك عزماً ابتدائياً من المرتبة الأولى على الأقل، فإذا كانا مستقلَّين بعضهما عن بعض، فعندئذ يكون لدينا $\mathbf{cov}(X,Y)=0$ .

**البرهان:** بسبب استقلال المتغيِّرين العشوائيين $X$ و$Y$، وباستخدام خصائص التوقع الرياضياتي السابقة، والأخذ بالحسبان خصائص المتسلسلات العددية المتقاربة (لأنَّه من الممكن أن تكون المجاميع غير منتهية)، فإنَّه يمكننا كتابة الآتي:

$$\mathbf{cov}(X,Y)=\mathbf{E}\left[(X-\mathbf{E}X)\cdot(Y-\mathbf{E}Y)\right]=\sum_{i\in I,j\in J}(x_i-\mathbf{E}X)\cdot(y_j-\mathbf{E}Y)\,P\left(X=x_i,Y=y_j\right)$$

$$=\sum_{i\in I,j\in J}(x_i-\mathbf{E}X)\cdot(y_j-\mathbf{E}Y)\,P\left(X=x_i\right)P\left(Y=y_j\right)=\mathbf{RHS}$$

ولكن لدينا:

$$\mathbf{RHS}=\sum_{i\in I}(x_i-\mathbf{E}X)P\left(X=x_i\right)\cdot\sum_{j\in J}(y_j-\mathbf{E}Y)P\left(Y=y_j\right)$$

$$=\left[\mathbf{E}(X-\mathbf{E}X)\right]\cdot\left[\mathbf{E}(Y-\mathbf{E}Y)\right]=\left[\mathbf{E}X-\mathbf{E}(\mathbf{E}X)\right]\cdot\left[\mathbf{E}Y-\mathbf{E}(\mathbf{E}Y)\right]$$

$$=\left[\mathbf{E}X-\mathbf{E}X\right]\cdot\left[\mathbf{E}Y-\mathbf{E}Y\right]=0$$

ومن ثمَّ ينتج لدينا أنَّ $\mathbf{cov}(X,Y)=0$ وبهذا يتمّ البرهان.

**(٧, ٢, ١, ١٨) ملاحظات**

١- إنَّ عكس المبرهنة السابقة غير صحيح في الحالة العامّة، والمثال الآتي (٧, ٢, ١, ١٩) يوضِّح لنا ذلك.

٢- إذا كان كلّ من $X$ و$Y$ يملك توقُّعاً رياضياتياً، وكان $\mathbf{cov}(X,Y)=0$ ، فعندئذ يُقال إنَّ المتغيِّرين العشوائيين $X$ و$Y$ غير مرتبطين Uncorrelated، وهكذا ينتج لدينا من الملاحظة السابقة أنّه إذا كان $X$ و$Y$ غير مرتبطين فإنَّ ذلك لا يعني أنَّهما مستقلّان. أي إنَّه إذا كان $\mathbf{E}(XY)=\mathbf{E}X\cdot\mathbf{E}Y$ فإنَّ ذلك **لا يعني** أنَّ المتغيِّرين العشوائيين $X$ و$Y$ مستقلّين.

**(٧, ٢, ١, ١٩) مثال**

ليكن $X$ و$Y$ متغيِّرين عشوائيين بسيطين فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$ وبمجموعتي قيم $\mathbf{X}=\{-1,2\}$ و $\mathbf{Y}=\{-2,-1,1,2\}$ على الترتيب، ويحقِّقان البندين الآتيين:

**a)** $P\left(X=-1,Y=+1\right)=P\left(X=-1,Y=-1\right)=\dfrac{1}{3}$

**b)** $P\left(X=2,Y=2\right)=P\left(X=2,Y=-2\right)=\dfrac{1}{6}$

لاحظ أنَّ باقي الاحتمالات الممكنة لـ $X$ و$Y$ معاً تساوي الصفر، وعلاوة على ذلك نجد ما يلي:

$$P\left(X=2\right)=P\left(X=2,Y=2\right)+P\left(X=2,Y=-2\right)=\frac{1}{6}+\frac{1}{6}=\frac{1}{3}$$

$$P\left(Y=2\right)=P\left(X=2,Y=2\right)=\frac{1}{6}$$

$$\frac{1}{6}=P\left(X=2,Y=2\right)\neq P\left(X=2\right)\cdot P\left(Y=2\right)=\frac{1}{18}$$

أي إنَّ $X$ و$Y$ ليسا مستقلَّين، وعلى الرغم من ذلك لدينا:

$$\mathbf{E}X=\sum_{\substack{x=2,\,x=-1\\ y=2,\,y=-2\\ y=1,\,y=-1}} x\ P(X=x)=2\left(\frac{1}{6}+\frac{1}{6}\right)+(-1)\left(\frac{1}{3}+\frac{1}{3}\right)=0$$

وبالمثل نجد $\mathbf{E}\,Y=0$ أيضاً، ومن ثَمَّ يصبح لدينا أنَّ $\mathbf{cov}(X,Y)=\mathbf{E}\,(XY)$، ولكن لدينا:

$$\mathbf{E}(XY)=\sum_{\substack{x=2,\,x=-1\\ y=2,\,y=-2\\ y=1,\,y=-1}}(x\cdot y)\,P(X=x\ ,\ Y=y)=4\left(\frac{1}{6}\right)-4\left(\frac{1}{6}\right)+1\left(\frac{1}{3}\right)-1\left(\frac{1}{3}\right)=0$$

ومنه ينتج لدينا أنَّ $\mathbf{cov}(X,Y)=0$، وهكذا أمكننا إيجاد متغيِّرين عشوائيين $X$ و$Y$ غير مستقلَّين وتغايرهما معدوماً.

### (٢٠ , ١ , ٢ , ٧) ملاحظات

١- في الحالة الخاصّة عندما يكون متغيِّران عشوائيان $X$ و$Y$ مستقلَّين، فإنَّه يصبح للعلاقة [7,15-a] العرض الآتي:

$$\mathbf{E}\,(XY)=\mathbf{E}X\cdot\mathbf{E}Y \qquad [7,15\text{-b}]$$

وذلك لأنَّه بسبب استقلال $X$ عن $Y$، وباستخدام خصائص المتسلسلات العددية المتقاربة يُمكننا أن نكتب الآتي:

$$\begin{aligned}\mathbf{E}(XY)&=\sum_{i\in I\ ,\ j\in J}(x_i\cdot y_j)\,P(X=x_i\ ,\ Y=y_j)=\sum_{i\in I\ ,\ j\in J}(x_i\cdot y_j)\cdot P(X=x_i)\cdot P(Y=y_j)\\ &=\sum_{i\in I}x_i\,P(X=x_i)\cdot\sum_{j\in J}y_j\,P(Y=y_j)=\mathbf{E}X\cdot\mathbf{E}Y\end{aligned}$$

نشير هنا إلى أنَّ هذه النتيجة غير قابلة للعكس، أي إنَّه إذا كانت العلاقة [7,15-b] مُحقَّقةً من أجل متغيِّرين عشوائيين $X$ و$Y$ فإنَّ ذلك لا يعني بالضرورة أنَّهما مستقلَّان بعضهما عن بعض، والمثال السابق (١٩ , ١ , ٢ , ٧) وضَّح لنا ذلك، حيث كان $X$ و$Y$ متغيِّرين عشوائيين غير مستقلَّين ومع ذلك كانت العلاقة [7,15-b] مُحقَّقة.

٢- إنَّ المقدار $\mathbf{E}(XY)$ يُدعى **العزم الارتباطي** بين $X$ و$Y$.

٣- بالرجوع إلى المثال (٩ , ١ , ٢ , ٧) حيث كان لدينا متغيِّرين عشوائيين $X$ و$Y$ يملكان توقعاً رياضياتياً مع $\mathbf{E}X=3.5$، و $\mathbf{E}Y=10.5$، فإنَّنا نجد:

$$\mathbf{E}\frac{X}{Y}=\sum_{i\in I,\,j\in J}\frac{x_i}{y_j}\cdot P\left(X=x_i,Y=y_j\right)=\frac{1}{6}\sum_{k=1}^{6}\frac{k}{3k}=\frac{1}{6}\sum_{k=1}^{6}\frac{1}{3}=\frac{1}{3}$$

ومنه يكون:

$$\frac{1}{3}=\mathbf{E}\frac{X}{Y}=\frac{\mathbf{E}X}{\mathbf{E}Y}=\frac{3.5}{10.5}=\frac{1}{3}$$

أي إنَّه من أجل ذلك المثال لدينا $\mathbf{E}\frac{X}{Y}=\frac{\mathbf{E}X}{\mathbf{E}Y}$، ولكن يجب الانتباه إلى أنَّه إذا كانا متغيِّرين عشوائيين $X$ و$Y$ يملكان توقُّعاً رياضياتياً، وبفرض أنَّ $P\left(Y\neq 0\right)=1$ (ومن ثَمَّ $\mathbf{E}Y\neq 0$) فعندئذ وجود هذا التوقُّع الرياضياتي لكلّ من $X$ و$Y$ لا يعني بالضرورة وجود التوقُّع الرياضياتي لـ $\frac{X}{Y}$، ومثال على ذلك إذا كان $X_1$ و$X_2$ متغيِّرين عشوائيين مستقلَّين، وكان ل $X_1$ توزيع طبيعي $\mathbf{N}(0,1)$ ولـ $X_2$ توزيع كاي مربَّع $\chi^2(1)$، فإنَّ كلّ منهما يملك توقعاً رياضياتياً حيث لدينا $\mathbf{E}X_1=0$ و $\mathbf{E}X_2=1$ (من أجل الاطلاع على قيم

التوقعات الرياضية للمتغيِّرات العشوائية الشهيرة التي سنوردها هنا ولاحقًا تجدها في جدول التوزيعات الاحتمالية الشهيرة في آخر هذا الكتاب)، فلو أخذنا $Y := \sqrt{X_2}$، فعندئذ يُبرهن على أنَّ للمتغيِّر العشوائي $Z = \frac{X_1}{Y}$ توزيع كوشي $\mathbf{C}(0,1)$ (انظر [65])، وهذا التوزيع لا يملك عزوماً ابتدائيةً ولا من أيّة مرتبة كما بيّنا ذلك سابقاً.

وأكثر من ذلك حتى في حال وجود $\mathbf{E}\frac{X}{Y}$ فإنَّه ليس بالضرورة أنَّ يكون $\mathbf{E}\frac{X}{Y}$ يساوي $\frac{\mathbf{E}X}{\mathbf{E}Y}$ ، وذلك لأنَّه في الحالة العامّة لدينا (انظر الملاحظ التالية) $\mathbf{E}\frac{X}{Y} \neq \frac{\mathbf{E}X}{\mathbf{E}Y}$، فلو أخذنا على سبيل المثال $X$ و$Y$ متغيِّرين عشوائيين مستقلّين بحيث إنَّ $X \sim \chi^2(n)$ مع $3 \le n$ و $Y \sim \chi^2(m)$ فيكون توقعهما الرياضياتي $\mathbf{E}X = n$ و $\mathbf{E}Y = m$ على الترتيب، علماً أنَّ $n$ و$m$ أعداد صحيحة موجبة، فلو أخذنا الآن $Z = \frac{X/n}{Y/m} = \frac{mX}{nY}$ فعندئذ يكون للمتغيِّر العشوائي $Z$ توزيع فيشر $\mathbf{F}(m;n)$ حيث لدينا $\mathbf{E}Z = \frac{n}{n-2}$ (انظر [65])، في حين أنَّ:

$$\frac{\mathbf{E}(mX)}{\mathbf{E}(nY)} = \frac{m \cdot \mathbf{E}X}{n \cdot \mathbf{E}Y} = \frac{m \cdot n}{n \cdot m} = 1 \neq \frac{n}{n-2} \quad for\ n \ge 3$$

٤- إذا كان كلّ من $X$ و$Y$ يملك توقُّعاً رياضياتياً، وبفرض أنَّ $P(Y \neq 0) = 1$، وأنَّ $g : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ دالَّة قيوسة بالنسبة إلى $\Re\times\Re$ و $\Re$، فعندئذ الصيغة التقريبية لحساب التوقُّع الرياضياتي لـ $g(X,Y)$ تُعطى بالعلاقة الآتية:

$$\begin{aligned}
\mathbf{E}\left[g(X,Y)\right] \approx g(\mathbf{E}X,\mathbf{E}Y) &+ \frac{1}{2}\,\mathbf{var}X\, \frac{\partial_2}{\partial x^2} g(x,y)\Big|_{x=\mathbf{E}X,\,y=\mathbf{E}Y} \\
&+ \frac{1}{2}\,\mathbf{var}Y\, \frac{\partial_2}{\partial y^2} g(x,y)\Big|_{x=\mathbf{E}X,\,y=\mathbf{E}Y} \\
&+ \mathbf{cov}(X,Y)\frac{\partial_2}{\partial x\,\partial y} g(x,y)\Big|_{x=\mathbf{E}X,\,y=\mathbf{E}Y}
\end{aligned}$$

التي ينتج عنها أنّه في حال أنَّ $P(Y \neq 0) = 1$ فإنَّه يكون لدينا:

$$\mathbf{E}\frac{X}{Y} \approx \frac{\mathbf{E}X}{\mathbf{E}Y} - \frac{1}{(\mathbf{E}Y)^2}\,\mathbf{cov}(X,Y) + \frac{\mathbf{E}X}{(\mathbf{E}Y)^3}\,\mathbf{var}Y \qquad [7,16]$$

٥- من أجل أي متغيِّر عشوائي غير سالب $X$ فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$ توجد متتالية متزايدة $(X_n)_{n\in\mathbb{N}}$ من المتغيِّرات العشوائية البسيطة فوق $[\Omega,\mathcal{A},P]$ بحيث إنَّه من أجل كلّ $\omega \in \Omega$ تكون العلاقة $X(\omega) = \lim_{n\to\infty} X_n(\omega)$ مُحقَّقةً، وحينئذ يُحسب التوقع الرياضياتي لـ $X$ من خلال العلاقة الآتية:

$$\mathbf{E}X = \lim_{n\to\infty} \mathbf{E}X_n \qquad [7,17]$$

٦- من المعلوم أنَّ كلّ متغيِّر عشـوائي كيفي $X$ فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$ يمكن كتابته على شكل فرق متغيِّرين عشوائيين غير سالبين من خلال العلاقة $X = X^+ - X^-$، علماً أنَّه من أجل كلّ $\omega \in \Omega$ لدينا:

$$X^+(\omega) := \mathbf{max}\{X(\omega), 0\} \qquad \& \qquad X^-(\omega) := -\mathbf{min}\{X(\omega), 0\}$$

فإذا كان $\mathbf{E}X^+ < +\infty$ أو $\mathbf{E}X^- < +\infty$، فعندئذ يكون التوقع الرياضياتي لـ $X$ مُعرَّفاً ويُعطى بالعلاقة الآتية:

$$\mathbf{E}X = \mathbf{E}X^+ - \mathbf{E}X^- \qquad [7,18]$$

ولكي يملك متغيِّر عشوائي كيفيّ $X$ توقعاً رياضياتياً يجب أن يملك كلّ من $X^+$ و $X^-$ توقعاً رياضياتياً أيضاً (أي إنَّ $\mathbf{E}X^+ < +\infty$ و $\mathbf{E}X^- < +\infty$)، وحينئذ يُحسب التوقع الرياضياتي لـ $X$ بالعلاقة [7,17]، ويكون لدينا:

$$\mathbf{E}X = \mathbf{E}X^+ - \mathbf{E}X^- < +\infty$$

توضِّح لنا المبرهنة الآتية إحدى الطرائق التي تساعدنا في تقدير قيمة التوقع الرياضياتي لمتغيِّر عشوائي غير سالب.

### (٧, ٢, ١, ٢١) مبرهنة (متباينة ماركوف Markov inequality)

إذا كان $X$ متغيِّراً عشوائياً غير سالب ويملك توقُّعاً رياضياتياً، فعندئذ من أجل كلّ $\varepsilon > 0$ (عدد حقيقي يمكن جعله صغيراً بالقدر الذي نريد) تكون المتباينة الآتية مُحقَّقة:

$$\mathbf{E}X \geq \varepsilon \cdot P(X \geq \varepsilon) \qquad [7,19\text{-a}]$$

وهذه المتباينة تُدعى متباينة ماركوف نسبةً إلى الرياضياتي الروسي ماركوف (1856-1922) Andrey Andreyevich Markov.

البرهان: على سبيل التبسيط سنثبت صحة هذه المتباينة من أجل متغيِّر عشوائي متقطِّع. لذلك سنأخذ $\varepsilon > 0$ ، فعندئذ يمكننا أن نكتب الآتي:

$$\begin{aligned}\mathbf{E}X &= \sum_{i\in I} x_i P(X = x_i) = \sum_{i\in I,\, x_i<\varepsilon} x_i P(X = x_i) + \sum_{i\in I,\, x_i\geq\varepsilon} x_i \cdot P(X = x_i)\\ &\geq \sum_{i\in I,\, x_i\geq\varepsilon} x_i P(X = x_i) \geq \varepsilon \sum_{i\in I,\, x_i\geq\varepsilon} P(X = x_i) = \varepsilon \cdot P(X \geq \varepsilon)\end{aligned}$$

وبهذا يتمّ البرهان.

### (٧, ٢, ١, ٢٢) ملاحظة

إنَّ المتباينة السابقة [7,19-a] تكافئ المتباينة الآتية:

$$F_X(\varepsilon) \geq 1 - \frac{\mathbf{E}X}{\varepsilon} \qquad [7,19\text{-b}]$$

التي تُمكِّننا من تقدير قيمة دالَّة التوزيع الاحتمالية لمتغيِّر عشوائي غير سالب $X$ لدى مُعرَّفة قيمته المتوقَّعة، أو تقدير التوقُّع الرياضياتي لمتغيِّر عشوائي غير سالب $X$ لدى مَعْرفة دالَّة توزيعه الاحتمالية.

تقدِّم لنا المبرهنة الآتية إحدى المتباينات المهمّة في مجالي العشوائيات والرياضيات التي تعود إلى الرياضياتي الروسي ليابانوف (1857-1918) Aleksandr Mikhailovich Lyapunov.

### (٧, ٢, ١, ٢٣) مبرهنة (متباينة ليابانوف Lyapunov Inequality)

إذا كان $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega, \mathcal{A}, P]$، ويملك توقُّعاً رياضياتياً حتى المرتبة $n$ ، فعندئذ من أجل أيّة قيم $\ell \in \mathbb{N}_n$ و $k \in \mathbb{N}_{n-1}$ مع $k < \ell$ تكون المتباينة الآتية (وتعرف باسم متباينة ليابانوف) مُحقَّقةً:

$$\left(\mathbf{E}\,|X|^k\right)^{\frac{1}{k}} \leq \left(\mathbf{E}\,|X|^\ell\right)^{\frac{1}{\ell}} \qquad [7,20]$$

البرهان: لنضع $m = \dfrac{\ell}{k}$ فيكون لدينا $1 < m$، ولنأخذ $Y = |X|^k$، فعندئذ بتطبيق متباينة جينسين على المتغيّر العشوائي $g(Y) = |Y|^m$ فنحصل على المتباينة $\left(\mathbf{E}|X|\right)^m \leq \mathbf{E}|X|^m$، ومن هذه المتباينة ينتج لدينا أنَّ:

$$\left(\mathbf{E}|X|^k\right)^{\frac{\ell}{k}} \le \mathbf{E}|X|^\ell$$

وبأخذ الجذر من المرتبة $\ell$ لطرفي المتباينة السابقة نحصل على المتباينة الآتية:

$$\left(\mathbf{E}|X|^k\right)^{\frac{1}{k}} \le \left(\mathbf{E}|X|^\ell\right)^{\frac{1}{\ell}}$$

وبهذا يتمّ البرهان.

**(٧,٢,١,٢) نتيجة**

تحت فرضيات المبرهنة السابقة نجد أنَّ:

$$\mathbf{E}|X| \le \left(\mathbf{E}|X|^2\right)^{\frac{1}{2}} \le \left(\mathbf{E}|X|^3\right)^{\frac{1}{3}} \le \dots \le \left(\mathbf{E}|X|^n\right)^{\frac{1}{n}}$$

والتي ينتج عنها أنَّه من أجل $n > k$ عدد صحيح سيكون لدينا ما يلي:

$$|\mathbf{E}X|^k \le \left(\mathbf{E}|X|\right)^k \le \mathbf{E}|X|^k$$

## (٧,٢,٢) العزوم المركزية لمتغيِّر عشوائي Centered Moments of Random Variables

من التطبيقات المهمّة للتوقُّع الرياضياتي لدالَّة قيوسة $g$ ما يُعرف باسم **العزوم حول نقطة** $c$ من $\mathbb{R}$ لمتغيِّر عشوائي التي سنبدأها بالتعريف الآتي.

**(٧,٢,٢,١) تعريف (العزم والعزم المطلق لمتغيِّر عشوائي حول نقطة)**

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$، و $2 \le k$ عدداً صحيحاً، و $c \in \mathbb{R}$ كيفياً ولكن مثَّبتاً، فعندئذٍ يُسمّى المقدار:

a) $\sigma_X^{(k)}(c) := \mathbf{E}(X - c)^k$ [7,21-a]

**العزم** حول $c$ من المرتبة $k$ للمتغيِّر العشـوائي $X$،

b) $\beta_X^{(k)}(c) := \mathbf{E}|X - c|^k$ [7,21-b]

**العزم المطلق** حول $c$ من المرتبة $k$ للمتغيِّر العشـوائي $X$.

الآن، يقال إنَّ متغيِّراً عشوائياً $X$ يملك عزماً حول $c$ من المرتبة $k$ إذا كان $\mathbf{E}|X - c|^k < +\infty$ ، أي إذا كان العزم المطلق حول $c$ من المرتبة $k$ للمتغيِّر العشوائي $X$ موجوداً.

**(٧,٢,٢,٢) ملاحظات**

١- في حال عدم وجود التباس (كأن تكون الدراسة متعلّقة بمتغيِّر عشوائي واحد $X$) سنكتب (على سبيل التبسيط) $\sigma^{(k)}(c)$ بدلاً من $\sigma_X^{(k)}(c)$، وكذلك $\beta^{(k)}(c)$ عوضاً عن $\beta_X^{(k)}(c)$.

٢- في الحالة الخاصّة إذا كان $c = 0$، فعندئذٍ يصبح لدينا:

$$\sigma^{(k)}(0) = \mathbf{E}X^k = \mu^{(k)} \qquad \& \qquad \beta^{(k)}(0) = \mathbf{E}|X|^k$$

٣- في الحالة الخاصّة عندما يكون $c = \mathbf{E}X$، فعندئذٍ سيكون $Y = X - \mathbf{E}X$ متغيِّراً عشـوائياً مركزياً، ولذلك فإنَّ:

$$\mathbf{E}Y^k = \mathbf{E}(X - \mathbf{E}X)^k$$

يُدعى **العزم المركزي من المرتبة** $k$ للمتغيِّر العشوائي $X$، وبالمثل يُدعى $\mathbf{E}|X-\mathbf{E}X|^k$ **العزم المركزي المطلق من المرتبة** $k$ للمتغيِّر العشوائي $X$.

٤- على سبيل التبسيط (وما لم يؤدِّ ذلك إلى التباس) سنكتب $\sigma^{(k)}$ بدلاً عن $\sigma^{(k)}(\mathbf{E}X)$، وكذلك $\beta^{(k)}$ بدلاً عن $\beta^{(k)}(\mathbf{E}X)$ أي إنَّه لدينا:

$$\sigma^{(k)} := \sigma^{(k)}(\mathbf{E}X) = \mathbf{E}\left(X-\mathbf{E}X\right)^k \qquad [7,22\text{-a}]$$

$$\beta^{(k)} := \beta^{(k)}(\mathbf{E}X) = \mathbf{E}\left|X-\mathbf{E}X\right|^k \qquad [7,22\text{-b}]$$

وقياسًا على ما سبق يُقال إنَّ متغيِّراً عشوائياً $X$ يملك عزماً مركزياً من المرتبة $k$ إذا كان العزم المركزي المطلق من المرتبة $k$ للمتغيِّر العشوائي $X$ موجوداً، أي إذا كان $\mathbf{E}|X-\mathbf{E}X|^k < +\infty$.

٥- في الحالة الخاصّة عندما يكون $k=2$ فإنَّ العزم المركزي من المرتبة الثانية للمتغيِّر العشوائي $X$ يُدعى **"التباين"** للمتغيِّر العشوائي $X$، ويُرمز له بـ $\mathbf{var}\,X$، أي إنَّه لدينا:

$$\sigma^{(2)} = \mathbf{E}\left(X-\mathbf{E}X\right)^2 \triangleq \mathbf{var}\,X \qquad [7,23]$$

٦- إنَّ الجذر التربيعي الموجب لتباين المتغيِّر العشوائي $X$ يُدعى **الانحراف المعياري** للمتغيِّر العشوائي $X$، ويُرمز له عادةً بـ $\sigma_X$، وفي حال عدم وجود التباس (كأن تكون الدراسة متعلِّقة بمتغيِّر عشوائي واحد فقط $X$) يُرمز له بـ $\sigma$ على سبيل التبسيط، أي إنَّ:

$$\sigma := +\sqrt{\mathbf{var}\,X} \qquad [7,24]$$

٧- من هذه العلاقة الأخيرة [7,24] ينتج لدينا $\sigma^2 = \mathbf{var}\,X$، وبما أنَّ $\mathbf{var}\,X = \sigma^{(2)}$، فلذلك يُكتب في هذه الحالة الخاصّة فقط $\sigma^{(2)} = \sigma^2$، وهذا يشير بدوره إلى أنَّ التباين هو مقدار غير سالب، حيث يُكتب مباشرة (في الكثير من الكتب والمراجع) أنَّ:

$$\sigma^2 = \mathbf{var}\,X = \mathbf{E}\left(X-\mathbf{E}X\right)^2 \qquad [7,25]$$

٨- يجب الانتباه هنا إلى أنَّ $\sigma^{(k)} \neq \sigma^k$ في الحالة العامّة؛ وذلك لأنَّ العزم المركزي $0 \le \sigma^k$ دوماً في حين أنَّ المقدار $\sigma^{(k)}$ قد يكون سالباً من أجل القيم الفردية لـ $k$، فعلى سبيل المثال لو أخذنا $X$ متغيِّراً عشوائياً متقطِّعاً ويأخذ جميع القيم الآتية بالاحتمال نفسه:

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 9 | 5 | 8 | 4 | 7 | 4 | 8 | 4 | 8 | 4 | 8 | 5 | 6 | 9 | 7 |

فعندئذ نجد أنَّ:

$$\mathbf{E}X = \sum_{i=1}^{15} x_i \cdot P(X=x_i) = \frac{1}{15}\sum_{i=1}^{15} x_i = 6.4$$

وكذلك نجد أنَّ $\sigma^2 = 3.44$، ومنه يكون $\sigma = 1.855$، ومن ثمَّ ينتج لدينا $\sigma^5 = 21.948$، في حين أنَّ:

$$\sigma^{(5)} = \sum_{i=1}^{15}(x_i-\mathbf{E}X)^5 \cdot P(X=x_i) = \frac{1}{15}\sum_{i=1}^{15}(x_i-\mathbf{E}X)^5 = -3.303$$

حيث نلاحظ أنَّ $\sigma^{(5)} \neq \sigma^5$.

٩- يُنظر إلى قيمة $\sigma^2 = \mathbf{var}\,X$ كمقياس لتبعثر قيم المتغيِّر العشوائي $X$ حول متوسط قيمه، ولكن يجب الانتباه إلى أنَّ القيمة الناتجة عنه تقرأ بالوحدة المربَّعة، ولذلك لا تستخدم في الجوانب التطبيقية، وإنَّما تستخدم قيمة الانحراف المعياري لـ $X$ كمقياس لتبعثر قيم $X$ حول متوسطها في الجوانب التطبيقية لأنَّ له وحدة القياس نفسها التي لقيم $X$.

١٠- يُبرهن (سنقبلها دون برهان) على أنَّه إذا كان متغيِّر عشوائي $X$ يملك عزماً ابتدائياً من المرتبة $k$ مع $k \in \mathbb{N}$ و $k \geq 2$، فإنَّه ستكون العلاقة الآتية مُحَقَّقةً:

$$\sigma^{(k)} = (-1)^{k-1}(k-1)\,\mu^{k} + \sum_{i=2}^{k} (-1)^{k-i} \binom{k}{i} \mu^{(i)}\,\mu^{k-i} \qquad [7,26]$$

وهذه العلاقة تفيدنا بأنَّ وجود العزم المركزي $\sigma^{(k)}$ لمتغيِّر عشوائي ينتج من وجوديّة العزوم الابتدائية $\mu^{(\ell)}$ لكل القيم $1 \leq \ell \leq k$ أيضاً.

١١- إنَّ تباين متغيِّر عشوائي $X$ يمكن أن يُحسب باستخدام العلاقة الآتية أيضاً:

$$\mathbf{var}\,X = \mathbf{E}X^{2} - (\mathbf{E}X)^{2} = \mu^{(2)} - \mu^{2} \qquad [7,27]$$

وهي تعرف باسم صيغة شتاينر Steiner Formula (سميت كذلك لتماثلها مع صيغة شتاينر - أو المبرهنة الأولى لـ هويغنز - في الميكانيكا).

سنقدِّم فيما يلي بعض خصائص التباين لمتغيِّر عشوائي دون إثبات.

### (٧,٢,٢,٣) خصائص التباين للمتغيِّرات العشوائية

ليكن $X$ و$Y$ متغيِّرين عشوائيين فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، وسنفترض أنَّ $\mathbf{E}X^{2} < +\infty$ وكذلك $\mathbf{E}Y^{2} < +\infty$، فعندئذ:

١- من أجل أي عدد حقيقي $c \in \mathbb{R}$ سيكون لدينا $\mathbf{var}\,X \leq \sigma^{(2)}(c)$، وهذا يعني أنَّ تباين متغيِّر عشوائي يُعدُّ أصغريًّا تحت عمليّة العزوم ذات المرتبة الثانية حول أي عدد $c \in \mathbb{R}$، وتكون إشارة المساواة في المتباينة السابقة مُحَقَّقةً إذا كان $P(X = c) = 1$، أي عندما يكون $X$ خاضعًا للتوزيع وحيد النقطة في $c$ فقط، ومن ثَمَّ سيكون لدينا في هذه الحالة الخاصّة (يترك برهانها تمرينًا للقارئ):

$$\mathbf{var}\,X = \sigma^{(2)}(c) = 0$$

٢- من أجل أي عددين $a$ و $b \in \mathbb{R}$ مع $a \neq 0$ تكون العلاقة الآتية مُحَقَّقة:

$$\mathbf{var}(aX + b) = a^{2}\,\mathbf{var}\,X \qquad [7,28]$$

ومن هذه الخاصيّة ينتج ما يلي:

$$\text{a)}\ \ \mathbf{var}(-X) = \mathbf{var}\,X \qquad \& \qquad \text{b)}\ \ \mathbf{var}\left(\frac{X}{\sqrt{\mathbf{var}\,X}}\right) = 1 \quad \textit{for}\ \ \mathbf{var}\,X > 0$$

٣- من تعريف التباين نستنتج أنَّ:

$$\mathbf{var}(X + Y) = \mathbf{var}\,X + \mathbf{var}\,Y + 2\cdot\mathbf{cov}(X,Y) \qquad [7,29\text{-a}]$$

وعندما يكون $X$ و$Y$ مستقلَّين فإنَّه سيكون $\mathbf{cov}(X,Y) = 0$، ومن ثَمَّ يصبح لدينا من أجل هذه الحالة الخاصّة:

$$\mathbf{var}(X + Y) = \mathbf{var}\,X + \mathbf{var}\,Y \qquad [7,29\text{-b}]$$

٤- لدينا:

$$\begin{aligned}\mathbf{var}(X\cdot Y) = {} & (\mathbf{E}Y)^{2}\,\mathbf{var}\,X + (\mathbf{E}X)^{2}\,\mathbf{var}\,Y + 2\mathbf{E}X\cdot\mathbf{E}Y\cdot\mathbf{cov}(X,Y) \\ & -\left[\mathbf{cov}(X,Y)\right]^{2} + \mathbf{E}\left[(X-\mathbf{E}X)^{2}\cdot(Y-\mathbf{E}Y)^{2}\right] \qquad [7,30\text{-a}] \\ & +2\mathbf{E}Y\cdot\mathbf{E}\left[(X-\mathbf{E}X)^{2}\cdot(Y-\mathbf{E}Y)\right] + 2\mathbf{E}X\cdot\mathbf{E}\left[(X-\mathbf{E}X)\cdot(Y-\mathbf{E}Y)^{2}\right]\end{aligned}$$

وفي حال كان $X$ و$Y$ مستقلَّان بعضهما عن بعض فإنَّه ينتج لدينا:

$$\mathbf{var}\ (X\cdot Y)=(\mathbf{E}Y)^2\ \mathbf{var}\ X+(\mathbf{E}X)^2\ \mathbf{var}\ Y+\mathbf{var}\ X\cdot\mathbf{var}\ Y \qquad [7,30\text{-}b]$$

**٥**- انطلاقاً من خواص التوقع الرياضياتي يتبيَّن لنا أنَّه تحت الفرضيات الواردة آنفًا وبفرض أنَّ $X\neq 0$ و $Y\neq 0$ باحتمال يساوي الواحد، فعندئذ ليس بالضرورة أنَّ يكون $\mathbf{var}\,\frac{X}{Y}$ موجوداً، وحتى في حال وجوده فقد يكون $\mathbf{var}\,\frac{X}{Y}\neq\frac{\mathbf{var}\,X}{\mathbf{var}\,Y}$، ولكنَّ الصيغة الآتية تُعطينا قيمة تقريبية لحساب $\mathbf{var}\,\frac{X}{Y}$ في حال وجودها:

$$\mathbf{var}\,\frac{X}{Y}\approx\left(\frac{\mathbf{E}X}{\mathbf{E}Y}\right)^2\cdot\left[\frac{\mathbf{var}\,X}{(\mathbf{E}X)^2}+\frac{\mathbf{var}\,Y}{(\mathbf{E}Y)^2}+2\frac{\mathbf{cov}(X,Y)}{\mathbf{E}X\cdot\mathbf{E}Y}\right] \qquad [7,31\text{-}a]$$

في الحقيقة إنَّ الصيغة السابقة تنتج عن صيغة أكثر عمومية لها العرض التالي:

$$\begin{aligned}\mathbf{var}\big(g(X,Y)\big)\approx\ &\mathbf{var}\,X\left[\frac{\partial}{\partial x}g(x,y)\Big|_{\mathbf{E}X,\,\mathbf{E}Y}\right]^2+\mathbf{var}\,Y\left[\frac{\partial}{\partial x}g(x,y)\Big|_{\mathbf{E}X,\,\mathbf{E}Y}\right]^2\\&+2\,\mathbf{cov}(X,Y)\left[\frac{\partial}{\partial x}g(x,y)\Big|_{\mathbf{E}X,\,\mathbf{E}Y}\ \frac{\partial}{\partial x}g(x,y)\Big|_{\mathbf{E}X,\,\mathbf{E}Y}\right]\end{aligned} \qquad [7,31\text{-}b]$$

علماً أنَّ $g$ دالَّة معرفة على $\mathbb{R}^2$ وقيوسة بالنسبة إلى $\mathfrak{R}^2$.

### (٧, ٢, ٢, ٤) ملاحظات

١- تجدر الإشارة هنا إلى أنَّ كلّ متغيِّر عشوائي تباينه يساوي الواحد يُدعى متغيِّراً عشوائياً **مُسْتَنظَماً** Normalized random variable، ولذلك فإنَّ عمليّة الانتقال من متغيِّر عشـوائي $X$ إلى متغيِّر عشوائي $Y$ مُعطى من خلال العلاقة الآتية:

$$Y=\frac{X}{\sqrt{\mathbf{var}\,X}}\qquad for\quad \mathbf{var}\,X>0 \qquad [7,32]$$

تُدعى عمليّة **استنظام** Normalization للمتغيِّر العشوائي $X$.

٢- إذا كان لمتغيِّر عشوائي $X$ توقُّع رياضياتي معدوم وتباين يسـاوي الواحد فعندئذ يُدعى $X$ متغيِّراً عشـوائياً **معيارياً (أو قياسياً)** Standard random variable، ومن ثمَّ فإنَّ عمليّة تحويل متغيِّر عشوائي $X$ إلى متغيِّر عشوائي مُعطى بالعلاقة الآتية:

$$Z=\frac{X-\mathbf{E}X}{\sqrt{\mathbf{var}\,X}}\qquad for\quad \mathbf{var}\,X>0 \qquad [7,33]$$

تُدعى عملية **اسْتعيّار (أو استقيّاس)** Standardization للمتغيِّر العشـوائي $X$.

قد نكون في بعض الأحيان بحاجة إلى حساب قيمة العزوم الابتدائية (ومن ثمَّ التباين ايضاً) لمتغيِّر عشوائي $X$، ووفقاً لما سبق قد يكون من غير الممكن في متابعة تنفيذ العمليات الحسابية لدى مرحلة من مراحل الحساب (حيث يكون استخدام الصيغ السابقة صعب جداً إن لم يكن غير ذي جدوى. انظر المثال /٢/ من الأمثلة القادمة (٧, ٢, ٤)). لذلك طوِّرت طريقة أخرى لحساب العزوم بحيث تساعد على تجاوز تلك الصعوبات في الحساب. هذه الطريقة عُرِفت باسم العزوم العَامليّة، وتقدِّمها لنا الفقرة التالية.

### (٧, ٢, ٣) العزم العَامِلي (أو الضَربِّي) لمتغيِّر عشوائي Factorial Moments of Random Variables

كما اوضحنا أعلاه فإنَّ للعزوم العامليّة دوراً فعَّالاً في حساب العزوم الابتدائية لمتغيرات عشوائية، والتعريف الآتي يقدِّم لنا هذا المفهوم.

**(٧ , ٢ , ٣ , ١) تعريف (العزم العاملي (أو الضربي) لمتغيِّر عشوائي)**

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، و $2 \le k$ عدداً صحيحاً، ولنأخذ المتغيّر العشوائي:

$$Y = g(X) := X\cdot(X-1)\cdot(X-2)\cdot\ldots\cdot(X-k+1)$$

فعندئذٍ يُدعى المقدار الآتي:

$$\mathbf{E}Y = \mathbf{E}[g(X)] = \mathbf{E}\left[X\cdot(X-1)\cdot(X-2)\cdot\ldots\cdot(X-k+1)\right]$$

**العزم العاملي (أو الضربي) من المرتبة $k$** للمتغيِّر العشوائي $X$ The $k^{\text{th}}$ Factorial Moment **of** $X$، وسنرمز له بـ $\mathcal{F}_X^{(k)}$ وفي حال عدم وجود التباس (أي إنَّنا نتعامل مع متغيِّر عشوائي واحد فقط) سنرمز له بـ $\mathcal{F}^{(k)}$ على سبيل التبسيط. أي إنَّه من أجل $2 \le k$ لدينا:

$$\mathcal{F}^{(k)} = \mathbf{E}\left[X\cdot(X-1)\cdot(X-2)\cdot\ldots\cdot(X-k+1)\right] \qquad [7,34]$$

**(٧ , ٢ , ٣ , ٢) ملاحظة**

إذا كان متغيِّراً عشوائياً $X$ يملك عزماً ابتدائياً من المرتبة الثانية، فإنَّه يمكن حساب التباين لـ $X$ باستخدام العزوم الابتدائية والعامليّة أيضاً كما في العلاقة الآتية:

$$\mathbf{var}\,X = \mathbf{E}[X(X-1)] + \mathbf{E}X\,(1-\mathbf{E}X) = \mathcal{F}^{(2)} + \mu(1-\mu) \qquad [7,35]$$

**(٧ , ٢ , ٤) أمثلة (حساب العزوم لبعض المتغيِّرات العشوائية الشهيرة)**

١ – لنأخذ تجربة إلقاء حجر نرد متوازن لمَّرة واحدة فقط، وليكن $X$ متغيِّراً عشوائياً فوق الفضاء الاحتمالي لهذه التجربة العشوائية مُعرَّفاً من خلال العلاقة الآتية:

$$X(\omega) = \begin{cases} -1 & for \quad \omega \ge 5 \\ 0 & for \quad \omega = 3,4 \\ +1 & for \quad \omega \le 2 \end{cases}$$

ولنقم بحساب القيمة الوسطى (أو المتوسط) لهذا المتغيِّر العشـوائي، ومن ثمَّ حساب التباين والانحراف المعياري لهذا المتغيّر العشوائي.

الحـل: لدينا تجربة إلقاء حجر نرد متوازن لمَّرة واحدة فقط، فعندئذ الفضاء الاحتمالي لهذه التجربة العشوائية هو $[\Omega,\mathscr{A},P]$ علماً أنَّ $\Omega = \{1,2,3,4,5,6\}$، وسنأخذ الـ σ – **جبر** فوق $\Omega$ هو $\mathscr{A} = 2^{\Omega}$، وأمّا الدالَّة الاحتمالية $P$ فيمكننا أخذها كما في العلاقة الآتية من أجل أي حادث $A \in \mathscr{A}$:

$$P(A) = \sum_{i:\,\omega_i \in A} Q(\{\omega_i\}) \qquad ; \ \mathbb{Q}(\{\omega_i\}) = p_i = \frac{1}{6} \ , \ i \in \mathbb{N}_6$$

وذلك لأنَّ حجر النرد متوازن. أمّا لحساب التباين للمتغيِّر العشوائي المُعطى فسوف نستخدم علاقة شتاينر [7,27]، وهذا يتطلب منّا حساب $\mathbf{E}X$ أولاً، وبما أنَّ المتغيّر العشوائي المُعطى متقطّعاً فإنَّ $\mathbf{E}X = \sum_{i\in I} x_i \cdot P(X = x_i)$، وبسبب توازن حجر النرد سيكون لدينا قيم الاحتمالات الكتلية هي: $P(X=-1)=\frac{1}{3}$، $P(X=0)=\frac{1}{3}$ و $P(X=1)=\frac{1}{3}$، وهذا يعني أنَّ المتغيِّر العشوائي $X$ يخضع للتوزيع المنتظم المتقطّع بمَعْلمة $n=3$، ومن ثمَّ يكون لدينا:

$$\mathbf{E}X = \sum_{i\in I} x_i \cdot P(X = x_i) = -1\times\frac{1}{3} + 0\times\frac{1}{3} + 1\times\frac{1}{3} = 0$$

أي إنَّ القيمة الوسطى (أو المتوسط) لهذا المتغيِّر العشوائي تساوي الصفر. أمّا العزم الابتدائي من المرتبة الثانية لـ $X$ فيُحسب بأسلوب مماثل، حيث لدينا:

$$\mathbf{E}\,X^2=\sum_{i\in I}x_i^2\cdot P(X=x_i)=(-1)^2\times\frac{1}{3}+(0)^2\times\frac{1}{3}+(1)^2\times\frac{1}{3}=\frac{2}{3}$$

وباستخدام علاقة شتاينر تكون قيمة التباين لهذا المتغيِّر العشوائي $X$ هي:

$$\mathbf{var}\,X=\mathbf{E}X^2-(\mathbf{E}X)^2=\frac{2}{3}-0=\frac{2}{3}=0.667$$

ومن ثَمّ يكون للانحراف المعياري القيمة الآتية:

$$\sigma=+\sqrt{\mathbf{var}\,X}=+\sqrt{0.667}=0.8167$$

٢- ليكن $X\sim\mathbf{U}(c)$ (خاضعًا للتوزيع الوحيد النقطة)، ولنقم بحساب التوقع الرياضياتي والتباين لهذا المتغيِّر العشوائي.

الحـل: نعلم أنَّ هذا المتغيِّر العشوائي متقطِّع بمجموعة قيم $\mathbf{X}=\{c\}$، وقانون توزيعه يُعطى بالعلاقة $P(X=c)=1$، فعندئذ يكون لدينا:

$$\mathbf{E}X=\sum_{i\in I}x_i\cdot P(X=x_i)=c\cdot P(X=c)=c$$

وبالمثل نجد:

$$\mathbf{E}X^2=\sum_{i\in I}x_i^2\cdot P(X=x_i)=c^2\cdot P(X=c)=c^2$$

ومنه باستخدام علاقة شتاينر يكون لدينا:

$$\mathbf{var}\,X=\mathbf{E}X^2-(\mathbf{E}X)^2=c^2-(c)^2=0$$

في الواقع إنَّ هذا النوع من المتغيِّرات العشوائية هو الوحيد الذي له تباين معدوم (يساوي الصفر) من بين جميع المتغيِّرات العشوائية التي تملك عزوماً ابتدائية من المرتبة الثانية. لا بل ويبرهن (سيترك كتمرين في آخر الفصل) على أنَّ كلّ متغيِّر عشوائي تباينه معدوم سيكون خاضعاً للتوزيع الوحيد النقطة.

لاحظ هنا أنّه لا يمكن إيجاد علاقة تربط بين التباين والتوقع الرياضياتي لـ $X$ إلاّ في الحالة الخاصَّة التي يكون فيها $c=0$، وذلك لأنَّ $\mathbf{var}X=0$ مهما تكن قيمة $\mathbf{E}X$.

٢- ليكن $X$ متغيِّراً عشوائياً خاضعاً للتوزيع الحدّاني $\mathbf{B}(n,p)$، ولنقم بحساب التوقع الرياضياتي والتباين لهذا المتغيِّر العشوائي.

الحـل: نعلم أنَّ هذا المتغيِّر العشوائي متقطِّع بمجموعة قيم $\mathbf{X}=\{0,1,2,...,n\}$ مع $n\in\mathbb{N}$ قيمة مثبَّتة، وقانون توزيعه يُعطى بالعلاقة الآتية:

$$P(X=k)=\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}$$

علماً أنَّ $0<p<1$، فعندئذ نجد أنَّ:

$$\mathbf{E}X = \sum_{k=0}^{n} k \cdot P\left(X = k\right) = \sum_{k=1}^{n} k \cdot \binom{n}{k} \cdot p^{k} \cdot (1-p)^{n-k} = \sum_{k=1}^{n} k \cdot \frac{n!}{k!(n-k)!} \cdot p^{k} \cdot (1-p)^{n-k}$$

$$= n\,p \sum_{k=1}^{n} \cdot \frac{(n-1)!}{(k-1)!\ [(n-1)-(k-1)]!} \cdot p^{k-1} \cdot (1-p)$$

$$\underset{\substack{n-1=m\\ k-1=\ell}}{=\!=\!=\!=\!=} n\,p \underbrace{\sum_{\ell=0}^{m} \cdot \frac{(m)!}{\ell!\ (m-\ell)!} \cdot p^{\ell} \cdot (1-p)^{m-\ell}}_{=1} = n \cdot p$$

وأمّا من أجل حساب تباين هذا المتغيّر العشوائي فإنّنا نلاحظ صعوبة حسابه باستخدام التعريف أو باستخدام علاقة شتاينر التي تعتمد على العزم من المرتبة الأولى، والثانية (حيث نلاحظ عدم جدوى استخدام تعريف العزم من المرتبة الثانية لحساب $\mathbf{E}X^2$)، ولذلك سنستخدم العزوم العامليّة من المرتبة الثانية، حيث لدينا:

$$\mathcal{F}^{(2)} = \mathbf{E}\left[X\ (X-1)\right] = \sum_{k=0}^{n} k\,(k-1) \cdot P\left(X = k\right)$$

$$= \sum_{k=2}^{n} k\,(k-1) \cdot \binom{n}{k} \cdot p^{k} \cdot (1-p)^{n-k} = \sum_{k=2}^{n} k\,(k-1) \cdot \frac{n!}{k!(n-k)!} \cdot p^{k} \cdot (1-p)^{n-k}$$

ومن ثمّ يمكننا أن نكتب:

$$\mathcal{F}^{(2)} = n\,(n-1)p^2 \sum_{k=2}^{n-2} \frac{(n-2)!}{(k-2)!\ [(n-2)-(k-2)]!} p^{k-2} (1-p)^{(n-2)-(k-2)}$$

$$\underset{\substack{n-2=m\\ k-2=\ell}}{=\!=\!=\!=\!=} n\,(n-1)p^2 \underbrace{\sum_{\ell=0}^{m} \cdot \frac{(m)!}{\ell!\ (m-\ell)!} \cdot p^{\ell} \cdot (1-p)^{m-\ell}}_{=1} = n\,(n-1)p^2$$

ومنه يصبح لدينا:

$$\mathbf{var}\,X = \mathbf{E}\left[X\ (X-1)\right] + \mathbf{E}X\ (1-\mathbf{E}X) = n\,(n-1)p^2 + n\,p\,(1-n\,p) = n\,p\,(1-p)$$

وهنا لاحظ هنا أنّه ثمّة علاقة تربط بين التباين والتوقع الرياضياتي لهذا المتغيّر العشوائي، حيث لدينا:

$$\mathbf{var}\,X = (1-p)\,\mathbf{E}X$$

وكتطبيق على هذا المثال سنفترض أنّ شخصاً يقوم بكتابة نص مكوّن من 5000 حرف، وأنّ احتمال خطأ هذا الشخص في كتابة حرف أثناء التنضيد يساوي 0.01، وأنّ كتابة كلّ حرف غير متأثرة بالحروف التي كتبت قبلها ولا تأثّر على كتابة الحروف التي تليها. عندئذ إذا كان $X$ متغيّراً عشوائياً راصداً لعدد الحروف التي كُتبت خطأً فإنّ $X$ سيكون خاضعاً للتوزيع الحدّاني $\mathbf{B}(5000,0.01)$، ومن ثمّ يكون متوسط عدد الحروف التي كُتبت خطأً في هذا النص يساوي:

$$\mathbf{E}X = n \cdot p = (5000) \cdot (0.01) = 50$$

بانحراف معياري يساوي:

$$\sigma = +\sqrt{\mathbf{var}\,X} = +\sqrt{n\,p\,(1-p)} = +\sqrt{(5000)(0.01)(1-0.01)} = 7.0356$$

أي إنّه من المتوقع أن يكون عدد الحروف التي كُتبت خطأً في هذا النص ما بين 43 و57 حرفاً تقريباً.

٣- ليكن $X$ متغيّراً عشوائياً خاضعاً للتوزيع الهندسي $\mathbf{G}(p)$، ولنقم بحساب قيمته المتوقعة ومن ثمّ تباينه.

الحـل: نعلم أنّ هذا المتغيّر العشوائي متقطّع بمجموعة قيم $\mathbf{X} = \mathbb{N}$، وقانون توزيع $P(X = k) = p\,(1-p)^{k-1}$ مع

$0 < p < 1$ ، ومن ثمَّ فإنَّ القيمة المتوقعة هذا المتغيِّر العشوائي هي:

$$\mathbf{E}X = \sum_{k=1}^{\infty} k \cdot P(X = k) = \sum_{k=1}^{\infty} k\ p(1-p)^{k-1} = p \sum_{k=1}^{\infty} k\ (1-p)^{k-1} = p \frac{d}{d\ (1-p)} \left( \sum_{k=1}^{\infty} (1-p)^k \right) = \frac{1}{p}$$

وهكذا نجد أنَّ متوسط قيم هذا المتغيِّر العشوائي يساوي مقلوب قيمة مَعْلمته $p$، وأمّا لحساب التباين فإنَّنا سنقوم أولاً بحساب العزم العاملي من المرتبة الثانية، حيث لدينا:

$$\mathcal{F}^{(2)} = \mathbf{E}\left[ X\ (X-1) \right] = \sum_{k=1}^{\infty} k\ (k-1) P(X = k) = \sum_{k=2}^{\infty} k\ (k-1) p\ (1-p)^{k-1}$$

$$= p\ (1-p) \frac{d\ 2}{d\ (1-p)^2} \left( \sum_{k=2}^{\infty} (1-p)^k \right) = \frac{2(1-p)}{p^2}$$

ومن ثمَّ يكون لدينا:

$$\mathbf{var}\,X = \frac{2(1-p)}{p^2} + \frac{1}{p} - \left( \frac{1}{p} \right)^2 = \frac{1-p}{p^2}$$

فلاحظ هنا أنَّ العلاقة التي تربط بين التباين والتوقع الرياضياتي لهذا المتغيِّر العشوائي هي $\mathbf{var}\,X = \frac{1-p}{p}\ \mathbf{E}X$ .

كتطبيق على هذا المثال لنأخذ المثال السابق، ولكن هنا سنفترض أنَّ الشخص سيتوقف عن الكتابة عند وقوع أول خطأ له في التنضيد، فعندئذ إذا كان $X$ متغيِّراً عشوائياً راصداً لعدد الحروف التي سيكتبها حتى يُخطئ لأول مرَّة، فإنَّ هذا المتغيِّر العشوائي سيكون خاضعاً للتوزيع الهندسي $\mathbf{G}(0.01)$ ، ومن ثمَّ يكون متوسط عدد الحروف التي سيكتبها حتى يُخطئ لأول مرَّة يساوي:

$$\mathbf{E}X = \frac{1}{p} = \frac{1}{0.01} = 100$$

وبانحراف معياري قدره:

$$\sigma = +\sqrt{\mathbf{var}\,X} = +\sqrt{\frac{1-p}{p^2}} = +\sqrt{\frac{1-0.01}{0.01^2}} = 99.499$$

لاحظ هنا كبَر قيمة الانحراف المعياري لهذا المتغيّر العشوائي، ومن ثمَّ من المتوقع أن يكون عدد الحروف التي سيكتبها حتى يُخطئ لأول مرَّة ما بين 1 و199 حرفاً تقريباً، وهذا يعني أنَّه من المتوقع ألّا يكتب ولا حرف صحيح ويتوقف عن الكتابة من أول حرف يكتبه، أو أن يصل عدد الحروف الصحيحة التي سيكتبها 199 حرفاً قبل توقفه عن الكتابة.

٤- ليكن $X$ متغيِّراً عشوائياً خاضعاً للتوزيع الأسي $\mathbf{EX}(\lambda)$ ولنقم بحساب التباين لهذا المتغيِّر العشوائي.

الحل: نعلم أنَّ هذا المتغيِّر العشوائي مستمرّ بدالَّة كثافة احتمالية:

$$f_X(x) = \lambda\,\mathbf{e}^{-\lambda x} \cdot \mathbf{I}_{[0,+\infty)}(x) \quad ; \forall x \in \mathbb{R}$$

علماً أنَّ $\lambda > 0$ ثابت، ولنقم بحساب العزوم الابتدائي من المرتبة الأولى والثانية لهذا المتغيِّر العشوائي، فنجدهما كما يلي:

$$\mathbf{E}X = \int_{-\infty}^{+\infty} x \cdot \lambda\,\mathbf{e}^{-\lambda x} \cdot \mathbf{I}_{[0,+\infty)}(x)\ dx = \int_{0}^{+\infty} \lambda\,x \cdot \mathbf{e}^{-\lambda x}\ dx = \frac{1}{\lambda} \int_{0}^{+\infty} y \cdot \mathbf{e}^{-y}\ dy = \frac{1}{\lambda} \cdot \Gamma(2) = \frac{1}{\lambda}$$

علماً أنَّ $\Gamma(x)$ ترمز إلى دالَّة غاما، حيث لدينا:

$$\int_{0}^{+\infty} x^{n}\, \mathbf{e}^{-x}\, dx = \Gamma(n+1) = n\,!$$

ولنحسب الآن $\mathbf{E}X^2$ فنجد ما يلي:

$$\begin{aligned}\mathbf{E}X^2 &= \int_{-\infty}^{+\infty} x^2 \cdot \lambda \mathbf{e}^{-\lambda x} \cdot \mathbf{I}_{[0,+\infty)}(x)\, dx \\ &= \int_{0}^{+\infty} \lambda x^2 \cdot \mathbf{e}^{-\lambda x}\, dx = \frac{1}{\lambda^2} \int_{0}^{+\infty} y^2 \cdot e^{-y}\, dy = \frac{1}{\lambda^2} \cdot \Gamma(3) = \frac{2}{\lambda^2}\end{aligned}$$

ومن ثمَّ يكون بحسب علاقة شتاينر لدينا:

$$\mathbf{var}\,X = \mathbf{E}X^2 - \left(\mathbf{E}X\right)^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2}$$

فلاحظ هنا أنَّ العلاقة التي تربط بين التباين والتوقع الرياضياتي لهذا المتغيِّر العشوائي هي:

$$\mathbf{var}\,X = (\mathbf{E}X)^2$$

كتطبيق على هذا المثال لنفترض أنّ لدينا جهازاً كهربائياً عمره يخضع للتو زيع الأسّي بمَعْلمة $\lambda = 0.01$ ساعة عمل (في الواقع توجد الكثير من الأجهزة الكهربائية التي عمرها الافتراضي يخضع للتوزيع الأسّي)، ولنقم بحساب متوسط عمر هذا الجهاز وتعيين أقل وأطول عمر مفترض له.

من أجل ذلك لنفترض أنَّ $X$ متغيِّراً عشوائياً راصداً لعمر هذا الجهاز، فعندئذ سيكون $X$ خاضعاً للتوزيع الأسي $\mathbf{EX}(0.01)$، ومن ثمَّ يكون متوسط عمر هذا الجهاز يساوي:

$$\mu = \mathbf{E}X = \frac{1}{\lambda} = \frac{1}{0.01} = 100$$

وهكذا نجد انَّ متوسط عمر هذا الجهاز يساوي 100 ساعة عمل، وأمّا لتعيين أقل وأطول عمر مفترض له فإنَّنا بحاجة لحساب الانحراف المعيار لـ $X$، فنجده بوساطة حساب التباين كما يلي:

$$\mathbf{var}\,X = \frac{1}{\lambda^2} = \frac{1}{0.0001} = 10000$$

ومن ثمَّ يكون $\sigma = +\sqrt{\mathbf{var}\,X} = 100$، وهذا يعني أنَّه من المتوقع أن يعمِّر الجهاز لمدّة تساوي $\mu \pm \sigma$، وهذا يعني أنَّه من المتوقع أن نجد الجهاز معطّلاً منذ بدء تشغيله كما يمكن له أن يعمل حتى 200 ساعة عمل.

٥- ليكن $X$ متغيِّراً عشوائياً خاضعاً للتوزيع الطبيعي $\mathbf{N}(\mu,\sigma)$، ولنقم بحساب التوقع الرياضياتي لهذا المتغيِّر العشوائي.

الحـل: نعلم أنَّ هذا المتغيِّر العشوائي مستمرّ بدالَّة كثافة احتمالية:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}}\, \mathbf{exp}\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad ; \forall x \in \mathbb{R}$$

علماً أنَّ $\mu \in \mathbb{R}$ و $\sigma > 0$ (ثوابت) هما معلمتي التوزيع، ولنقم أولاً بحساب العزم الابتدائي من المرتبة الأولى لهذا المتغيِّر العشوائي فنجد الآتي:

$$\mathbf{E}X=\int_{-\infty}^{\infty}x\,f_X(x)\;dx=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}x\cdot\mathbf{exp}\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx=\frac{\sqrt{2}\cdot\sigma}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{(x-\mu)+\mu}{\sqrt{2}\cdot\sigma}\;\mathbf{exp}\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx$$

$$=\underbrace{\frac{\sqrt{2}\cdot\sigma}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{(x-\mu)}{\sqrt{2}\cdot\sigma}\;\mathbf{exp}\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx}_{=0}+\mu\underbrace{\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}\;\mathbf{exp}\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]dx}_{=1}=\mu$$

كما يمكن إثبات (يترك كتمرين في آخر الفصل) أنَّ تباين هذا المتغيِّر العشوائي يساوي $\sigma^2$، أي إنَّ $\mathbf{var}X=\sigma^2$، فلاحظ هنا أنَّه لا توجد علاقة نظرية مُحدَّدة تربط بين قيمة التباين والتوقع الرياضياتي لهذا المتغيِّر العشوائي (ولهذا الكلام معنى سنذكره لاحقاً في نظرية المعاينة).

كتطبيق على هذا المثال لنفترض أنّه لدينا آلة لإنتاج نوع من القضبان المعدنية ذات قطر يساوي 12 مم، فإذا علمنا أنَّ قيم أقطار القضبان المنتجة من قبل هذه الآلة تخضع للتوزيع الطبيعي بمتوسط 12مم وانحراف معياري 0.5 مم (في الواقع يوجد الكثير من القياسات المعمليّة تخضع للتوزيع الطبيعي)، فإذا أخذنا قضيبًا ما من إنتاج هذه الآلة فما هو احتمال ألّا يقل قطره عن 12 مم؟

الحـل: من أجل ذلك لنفترض أنَّ $X$ متغيِّراً عشوائياً راصداً لقطر القضيب المنتج من هذه الآلة، فعندئذ سيكون $X$ خاضعاً للتوزيع الطبيعي $\mathbf{N}(12,0.5)$، ومن ثَّم يكون الاحتمال المطلوب هو $P(X\geq 12)$، ولحساب هذا الاحتمال لدينا:

$$P(X\geq 12)=1-P(X<12)=1-P\left(\frac{X-\mu}{\sigma}<\frac{12-\mu}{\sigma}\right)=1-P\left(Z<0\right)=1-0.5=0.5$$

حيث قمنا باستعّيار المتغيِّر العشوائي $X$ باستخدام التحويل $Z=\dfrac{X-12}{0.5}$، ولدينا من جدول التوزيع الطبيعي $P\left(Z<0\right)=0.5$.

### (٧, ٢, ٥) متباينات شهيرة ذات صلة بالعزوم

قد نكون في بعض الأحيان بحاجة إلى تقدير قيمة الحد الأدنى لقيمة التباين لمتغيِّر عشـوائي $X$. في هذا المجال تقدّم لنا المتباينة الآتية تقديراً جيداً لقيمة تباين متغيِّر عشـوائي $X$.

#### (٧, ٢, ٥, ١) مبرهنة (متباينة تشبيشيف Chebyshev inequality)

ليكن $X$ متغيِّراً عشـوائياً فوق فضاء احتمالي $\left[\Omega,\mathcal{A},P\right]$ مع $\mathbf{E}X^2<+\infty$، فعندئذ من أجل أي $\varepsilon>0$ تكون المتباينة الآتية مُحقّقة:

$$\mathbf{var}\,X\geq\varepsilon^2\cdot P\left((X-\mathbf{E}X)^2\geq\varepsilon^2\right)\qquad[7,36\text{-a}]$$

وهذه المتباينة تُعرف باسم متباينة تشيبيشيف.

البرهان: بملاحظة أنَّ المتغيِّر العشوائي $Y=(X-\mathbf{E}X)^2$ غير سالب ويملك توقُّعاً رياضياتياً لأنَّ $\mathbf{E}X^2<+\infty$، فإنَّه بحسب متباينة ماركوف [7,19-a] سيكون الآتي مُحقَّقاً من أجل أي $\varepsilon^2=\delta>0$:

$$\mathbf{var}\,X=\mathbf{E}Y\geq\delta\cdot P(Y\geq\delta)=\varepsilon^2\cdot P\left((X-\mathbf{E}X)^2\geq\varepsilon^2\right)$$

وبذلك يتمّ البرهان.

#### (٧, ٢, ٥, ٢) ملاحظات

١ – من المتباينة [7,36-a] يمكننا أن نكتب الآتي من أجل كلّ $x>0$:

$$\frac{\mathbf{var}\,X}{x}\geq 1-P\left((X-\mathbf{E}X)^2<x\right)=1-F_{(X-\mathbf{E}X)^2}(x)$$

والتي نجد منها:

$$F_{(X-\mathbf{E}X)^2}(x) \geq 1 - \frac{\mathbf{var}\,X}{x} \qquad [7,36\text{-b}]$$

وهذه العلاقة تساعدنا في تقدير قيمة دالّة التوزيع الاحتمالية للمتغيِّر العشوائي $Y = (X - \mathbf{E}X)^2$.

٢- تُعدّ متباينة تشبيشيف من أفضل المتباينات المستخدمة لتقدير قيمة تباين متغيّر عشوائي يملك عزوماً ابتدائية من المرتبة الثانية على الأقل.

٣- من متباينة كوشي- شفارتز (المبرهنة (٧, ٢, ١, ٨)) يمكننا الحصول على المتباينة الآتية:

$$\left[\mathbf{cov}\,(X\,,\,Y\,)\right]^2 \leq \mathbf{var}X \cdot \mathbf{var}Y \qquad [7,37\text{-a}]$$

ويتمّ ذلك من خلال تطبيق تلك المتباينة على المتغيِّرين العشوائيين $X = (U - \mathbf{E}U)$ و $Y = (T - \mathbf{E}T)$، علماً أنَّ $U$ و $T$ متغيِّرين عشوائيين فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ مع $\mathbf{E}U^2 < +\infty$ و $\mathbf{E}T^2 < +\infty$. أكثر من ذلك تُمكّننا المتباينة السابقة [7,37-a] من تقدير قيمة التباين لأحد المتغيِّرين العشوائيين $X$ أو $Y$ بدلالة قيمة تغايرهما والتباين للمتغيِّر الآخر، وذلك كما يلي:

$$\mathbf{var}Y \geq \frac{1}{\mathbf{var}\,X}\left[\mathbf{cov}\,(X\,,Y\,)\right]^2 \qquad [7,37\text{-b}]$$

### (٧, ٢, ٦) مُعامل الارتباط الخطي لمتغيِّرين عشوائيين Linear Correlation Coefficient of two R.V.

لقد رأينا في الفصل الثالث من هذا الكتاب أنّه من المقاييس المهمّة التي تعتمد على العزوم المركزية لمجموعتي بيانات هو الارتباط الخطي بين هاتين المجموعتين، حيث لاحظنا أنَّ البيانات تتولّد عادة من متغيِّرات، ولذلك فإنَّ هذا المفهوم ينسحب على المتغيِّرات العشوائية أيضاً، ولكن بفارق وحيد هو أنَّ القيم التي سنتعامل معها في هذا القطّاع هي قيم عشوائية حتماً. التعريف الآتي يقدِّم لنا كيفية حساب مُعامل الارتباط الخطي بين متغيِّرين عشوائيين.

#### (٧, ٢, ٦, ١) تعريف مُعامل الارتباط الخطي لمتغيِّرين عشوائيين

ليكن $X$ و $Y$ متغيِّرين عشوائيين فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، وكلّ منهما يملك عزماً ابتدائياً من المرتبة الثانية، فعندئذ يُرمز بـ $\rho_{X,Y}$ لمُعامل الارتباط الخطي بين المتغيِّرين العشوائيين $X$ و $Y$، ويُعرَّف من خلال العلاقة الآتية:

$$\rho_{X,Y} := \begin{cases} \dfrac{\mathbf{cov}(X\,,Y\,)}{\sqrt{\mathbf{var}\,X}\,\sqrt{\mathbf{var}Y}} & \textit{for } \mathbf{var}\,X > 0 \textit{ and } \mathbf{var}Y > 0 \\ 0 & \textit{otherwise} \end{cases} \qquad [7,38]$$

#### (٧, ٢, ٦, ١) ملاحظات

١- إذا كان $X$ مستقلاً عن $Y$، فعندئذ سيكون $\rho_{X,Y} = 0$، ولكنَّ عكس هذه النتيجة غير صحيح، بمعنى أنّه إذا كان $\rho_{X,Y} = 0$ فإنَّ ذلك **لا يعني** بالضرورة أن يكون $X$ و $Y$ مستقلّان بعضهما عن بعض (أوجد مثالاً مناسبا على ذلك).

٢- من الواضح أنَّ $-1 \leq \rho_{X,Y} \leq +1$، وأنَّ هذا الارتباط يكون إيجابياً إذا كان $\rho_{X,Y} > 0$، وعكسياً إذا كان $\rho_{X,Y} < 0$، وعندما يصبح $\rho_{X,Y} = \pm 1$ فعندئذ يكون المتغيِّرين العشوائيين $X$ و $Y$ مرتبطان مع بعضهما بعلاقة خطيّة، أي إنَّه يمكننا أن نكتب:

$$Y = a + b\,X$$

أوجد مثالاً مناسباً على كلّ حالة من هذه الحالات المذكورة آنفاً.

٣- بقيّة الخصائص لهذا المُعامل تناقش بشكل مماثل كما مرَّ معنا سابقاً (في الفصل الثالث) بخصوص مُعامل الارتباط الخطي.

## (٧,٣) الدالة المولدة الاحتمالية

**Probability Generating Function**

إنَّ الدوال المولَّدة الاحتمالية تقوم بدور مهمٍّ في تعيين الكثير من التوزيعات الاحتمالية للمتغيِّرات العشوائية، حتى إنَّ بعض التوزيعات الاحتمالية الشهيرة سميَّت وفقاً للنّموذج الرياضياتي للدالَّة المولَّدة الاحتمالية له، ومثال على ذلك التوزيع الهندسي وفوق الهندسي، وذلك لأنَّ الدالَّة المولَّدة الاحتمالية لمتغيِّر عشوائي خاضع للتوزيع الهندسي هي دالَّة هندسية في حين أنَّ الدالَّة المولَّدة الاحتمالية لمتغيِّر عشوائي خاضع للتوزيع فوق الهندسي هي دالَّة فوق هندسية.

**(٧,٣,١) تعريف (الدالَّة المولَّدة الاحتماليةلتغيِّر عشوائي)**

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، ولنعرِّف من أجل هذا المتغيِّر العشوائي دالَّة حقيقية $\mathbf{G}_X$ على $\mathbb{R}$ كما يلي:

$$\mathbf{G}_X : \mathbb{R} \longrightarrow \mathbb{R} \quad ; s \mapsto \mathbf{G}_X(s) := \mathbf{E}s^X \qquad [7,39\text{-}a]$$

فعندئذ تُدعى $\mathbf{G}_X$ **الدالَّة المولَّدة الاحتمالية** للمتغيِّر العشوائي $X$.

**(٧,٣,١,١) ملاحظات**

١ – إذا كان $X$ متغيِّراً عشوائياً متقطِّعاً مجموعة قيمه $\mathbf{X} = \{x_i \; ; i \in I\}$، فعندئذ يُصبح للدالَّة المولَّدة الاحتمالية لـ $X$ العرض الآتي:

$$\mathbf{G}_X(s) = \sum_{i \in I} s^{x_i} P(X = x_i) \qquad [7,39\text{-}b]$$

وفي حال كانت مجموعة قيمه $\mathbf{X} = \mathbb{N}$، فعندئذ تُكتب الدالَّة المولَّدة الاحتمالية لـ $X$ على النحو الآتي:

$$\mathbf{G}_X(s) = \sum_{k \in \mathbb{N}} s^k P(X = k) \qquad [7,39\text{-}c]$$

ولكي يملك المتغيِّر العشوائي $X$ دالَّة مولَّدة احتمالية يجب أن تكون أيّ من المتسلسلتين العدديتين السابقتين [7,39-b] و [7,39-c] متقاربةً مطلقاً، حيث نلاحظ أنَّ المتسلسلة العددية الأخيرة تكون متقاربة من أجل جميع قيم $s$ المحقِّقة للمتباينة $|s| \le 1$.

٢ – إذا كان $X$ متغيِّراً عشوائياً مستمراً بدالَّة كثافة احتمالية $f_X$، فعندئذ يُصبح للدالَّة المولَّدة الاحتمالية لـ $X$ العرض الآتي:

$$\mathbf{G}_X(s) = \int_{-\infty}^{\infty} s^x f_X(x)\, dx \qquad [7,39\text{-}d]$$

ولكي يملك متغيِّر عشوائي مستمرّ $X$ دالَّة مولَّدة احتمالية يجب أن يكون التكامل المعتل في العلاقة [7,39-d] متقارباً مطلقاً.

**(٧,٣,١,٢) أمثلة**

١ – لنأخذ $X \sim \mathbf{D}(c)$، فعندئذ يكون لدالَّته المولَّدة الاحتمالية العرض الآتي:

$$\mathbf{G}_X(s) = s^c \qquad ; \forall s \in \mathbb{R}$$

وذلك لأنَّه من أجل كلّ $s \in \mathbb{R}$ لدينا:

$$\mathbf{G}_X(s) = \mathbf{E}\left(s^X\right) = \sum_{k \in I} s^{x_k} P(X = x_k) = s^c \cdot P(X = c) = s^c$$

٢– لنأخذ $X \sim \mathbf{B}(n,p)$ ، فعندئذ يكون لدالّته المولّدة الاحتمالية العرض الآتي:

$$\mathbf{G}_X(s)=\left[p(s-1)+1\right]^n \qquad ;\forall s \in \mathbb{R}$$

وذلك لأنّه من أجل كلّ $s \in \mathbb{R}$ لدينا:

$$\mathbf{G}_X(s)=\mathbf{E}_X\left(s^X\right)=\sum_{i\in I} s^{x_i}\,P(X=x_i)=\sum_{k=0}^{n} s^k P(X=k)=\sum_{k=0}^{n} s^k\cdot\binom{n}{k}\cdot p^k\cdot(1-p)^{n-k}$$
$$=\sum_{k=0}^{n}\binom{n}{k}\cdot(s\,p)^k\cdot(1-p)^{n-k}=\left[sp+(1-p)\right]^n=\left[p(s-1)+1\right]^n$$

٣– لنأخذ $X \sim \mathbf{G}(p)$ ، فعندئذ يكون لدالّته المولّدة الاحتمالية العرض الآتي:

$$\mathbf{G}_X(s)=\frac{p\,s}{1-s(1-p)} \qquad ;\forall s \in \mathbb{R}$$

وذلك لأنّه من أجل كلّ $s \in \mathbb{R}$ لدينا:

$$\mathbf{G}_X(s)=\mathbf{E}(s^X)=\sum_{k=1}^{\infty} s^k\,P(X=k)=\sum_{k=1}^{\infty} s^k\,(1-p)^{k-1}\cdot p=s\,p\sum_{\ell=0}^{\infty}\left(s(1-p)\right)^{\ell}=\frac{p\,s}{1-s(1-p)}$$

حيث نلاحظ أنّ الطرف الأيمن من العلاقة الأخيرة هي صيغة لدالّة هندسية، ومن أجل الحالة الخاصّة $p=\frac{1}{2}$ (التساوي في النصيب للتجربة البرنولية) يصبح لدينا:

$$\mathbf{G}_X(s)=\sum_{k=1}^{\infty} s^k\,\frac{1}{2}\left(\frac{1}{2}\right)^{k-1}=\frac{s}{2-s} \qquad ;\forall s \in \mathbb{R}\setminus\{2\}$$

٤– لنأخذ $X \sim \mathbf{P}(\lambda)$ ، فعندئذ يكون لدالّته المولّدة الاحتمالية العرض الآتي:

$$\mathbf{G}_X(s)=\mathbf{exp}\left\{\lambda(s-1)\right\} \qquad ;\forall s \in \mathbb{R}$$

وذلك لأنّه من أجل كلّ $s \in \mathbb{R}$ لدينا:

$$\mathbf{G}_X(s)=\mathbf{E}\,s^X=\sum_{k=0}^{\infty} s^k\,P(X=k)=\sum_{k=0}^{\infty} s^k\,\frac{\lambda^k}{k\,!}\,\mathbf{e}^{-\lambda}=\mathbf{e}^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda\,s)^k}{k\,!}=\mathbf{e}^{-\lambda}\,\mathbf{e}^{\lambda s}=\mathbf{e}^{\lambda(s-1)}$$

٥– لنأخذ $X \sim \mathbf{U}\left([a,b]\right)$ ، فعندئذ يكون لدالّته المولّدة الاحتمالية العرض الآتي:

$$\mathbf{G}_X(s)=\frac{s^b-s^a}{(b-a)\,\mathbf{ln}\,s} \qquad ;\forall s \in \mathbb{R}$$

وذلك لأنّه من أجل كلّ $s \in \mathbb{R}$ لدينا:

$$\mathbf{G}_X(s)=\int_{-\infty}^{\infty} s^x f_X(x)\,dx=\frac{1}{b-a}\int_a^b s^x\,dx=\frac{1}{b-a}\left(\left.\frac{s^x}{\mathbf{ln}\,s}\right|_a^b\right)=\frac{s^b-s^a}{(b-a)\,\mathbf{ln}\,s}$$

### (٧ ، ٣ ، ٢) خصائص الدالّة المولّدة الاحتمالية

المبرهنة الآتية التي سنقبلها دون برهان تقدّم لنا بعض الخصائص المهمّة للدالّة المولّدة الاحتمالية التي تساعدنا في حساب العزوم للمتغيّرات العشوائية التي تملك عزوماً ابتدائية.

**(٧,٣,٢,١) مبرهنة**

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ بدالَّة مولِّدة احتمالية $\mathbf{G}_X$ ويملك عزماً ابتدائياً حتى المرتبة $k$ مع $k \in \mathbb{N}$ كيفي ولكن مثبَّت، فعندئذ سيكون لدينا ما يلي مُحَقَّقاً:

a) $\mathbf{G}_X(1)=1$

b) $\mathbf{G}_X(0)=P(X=0)$

c) $\mathbf{E}X=\mathbf{G}'_X(1)$

d) $\mathbf{E}\left[X\cdot(X-1)\cdot\ldots\cdot(X-k+1)\right]=\lim\limits_{s\uparrow 1}\mathbf{G}_X^{(k)}(s)$

علماً أنَّ $\mathbf{G}_X^{(k)}(s)$ هو المشتّق من المرتبة $k$ للدالَّة $\mathbf{G}_X$ عند النقطة $s$، ويُشار في بعض الأحيان إلى المقدار $\lim\limits_{s\uparrow 1}\mathbf{G}_X^{(k)}(s)$ بالرمز $\mathbf{G}_X^{(k)}(1)$ أي إنَّه يُكتب:

$$\mathbf{E}\left[X\cdot(X-1)\cdot\ldots\cdot(X-k+1)\right]=\mathbf{G}_X^{(k)}(1)$$

**(٧,٣,٢,٢) أمثلة**

١– لنأخذ $X \sim \mathbf{D}(c)$، فعندئذ يكون من أجل كلّ $s \in \mathbb{R}$ لدينا:

$$\mathbf{G}'_X(s)=c\cdot s^{c-1} \quad \& \quad \mathbf{G}''_X(s)=c(c-1)\cdot s^{c-2} \quad \& \quad \mathbf{G}'''_X(s)=c(c-1)(c-2)\cdot s^{c-3}$$

ومن ثمَّ يكون بحسب المبرهنة السابقة (٧,٣,٢,١) لدينا ما يلي:

$$\mathbf{E}X=\mathbf{G}'_X(1)=c\cdot s^{c-1}\Big|_{s=1} \Rightarrow \mathbf{G}'_X(1)=c$$

وكذلك:

$$\mathbf{E}\left[X\cdot(X-1)\right]=\mathbf{G}''_X(1)=c(c-1)\cdot s^{c-2}\Big|_{s=1}=c(c-1)$$

ومنه ينتج لدينا أنَّ:

$$\mathbf{E}X^2=\mathbf{E}\left[X\cdot(X-1)\right]+\mathbf{E}X=c(c-1)+c=c^2$$

وكذلك نجد:

$$\mathbf{E}X^3=\mathbf{E}\left[X\cdot(X-1)\cdot(X-2)\right]+3\mathbf{E}X^2-2\mathbf{E}X=c(c-1)(c-2)+3c^2-2c=c^3$$

وباستخدام صيغة شتاينر نجد أنَّ التباين لهذا المتغيِّر العشوائي يساوي الصفر، حيث لدينا:

$$\mathbf{var}X=\mathbf{E}X^2-(\mathbf{E}X)^2=\mu^{(2)}-\mu^2=c^2-c^2=0$$

وهي ذات النتيجة التي وصلنا إليها سابقاً بخصوص التباين لهذا النوع من المتغيِّرات العشوائية.

٢– لنأخذ $X \sim \mathbf{B}(n,p)$، فعندئذ يكون من أجل كلّ $s \in \mathbb{R}$ لدينا:

$$\mathbf{G}'_X(s)=\left(\left[p(s-1)+1\right]^n\right)'=n\,p\left[p(s-1)+1\right]^{n-1}$$

ومن ثمَّ يكون لدينا بحسب المبرهنة السابقة (٧,٣,٢,١) ما يلي:

$$\mathbf{E}X=\mathbf{G}'_X(1)=n\,p\left[p(s-1)+1\right]^{n-1}\Big|_{s=1}=n\,p$$

وهي ذات القيمة التي حصلنا عليها سابقاً ولكن هنا بطريقة أسهل وأسرع في الحساب. كذلك لدينا:

$$\mathbf{E}\left[X\cdot(X-1)\right]=\mathbf{G}''_X(1)=n(n-1)p^2\left[p(s-1)+1\right]^{n-2}\Big|_{s=1}=n(n-1)p^2$$

ومنه ينتج لدينا أنَّ:

$$\mathbf{E}X^2=\mathbf{E}\left[X\cdot(X-1)\right]+\mathbf{E}X=n(n-1)p^2+np=np\left[(n-1)p+1\right]=\left[(n-1)p+1\right]\cdot\mathbf{E}X$$

وبالمثل نجد:

$$\mathbf{E}\left[X(X-1)(X-2)\right]=\mathbf{G}'''_X(1)=n(n-1)(n-2)p^3\left[p(s-1)+1\right]^{n-3}\Big|_{s=1}=n(n-1)(n-2)p^3$$

ومن ثَمّ ينتج لدينا:

$$\begin{aligned}\mathbf{E}X^3&=\mathbf{E}\left[X\cdot(X-1)\cdot(X-2)\right]+3\mathbf{E}X^2-2\mathbf{E}X=n(n-1)(n-2)p^3+3n(n-1)p^2-2np\\&=np\left(n^2p^2-3np^2+2p^2+3np-3p-2\right)=\left[(n^2-3n+2)p^2+(3n-3)p-2\right]\cdot\mathbf{E}X\end{aligned}$$

٣– لنأخذ $X\sim\mathbf{G}(p)$، فعندئذ يكون من أجل كلّ $s\in\mathbb{R}$ لدينا:

$$\mathbf{G}'_X(s)=\left(\frac{ps}{1-s(1-p)}\right)'=\frac{p\left[1-s(1-p)\right]+(1-p)ps}{\left[1-s(1-p)\right]^2}$$

ومن ثَمّ يكون لدينا بحسب المبرهنة السابقة (١, ٢, ٣, ٧) ما يلي:

$$\mathbf{E}X=\mathbf{G}'_X(1)=\frac{p\left[1-s(1-p)\right]+(1-p)ps}{\left[1-s(1-p)\right]^2}\Bigg|_{s=1}=\frac{p^2+(1-p)p}{p^2}=\frac{1}{p}$$

وهي ذات القيمة التي حصلنا عليها سابقاً، ولكن هنا بطريقة أسهل، ودون التطرق لاشتقاق المتسلسلات الدالّية.

إنَّ نظرية الاحتمالات كثيراً ما تهتمّ بمجاميع لمتغيِّرات عشوائية، ولدراسة مثل هذه المجاميع نحتاج إلى طريقة مفيدة وفعَّالة لوصف توزيع هذه المجاميع. في الواقع، وكما ذكرنا آنفاً، أثبتت الدوال المولِّدة الاحتمالية أنَّها مصدر قوَّة في هذا المجال، فلو كان $X$ و$Y$ متغيِّرين عشوائيين متقطِّعين ومستقلَّين فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فعندئذ الدالَّة الاحتمالية لمجموعهما تُعطى بالعلاقة الآتية:

$$P\left(X+Y=u\right)=\sum_{x\in\mathbf{X}}P\left(X=x\right)\cdot P\left(Y=u-x\right)$$

والذي يُلزمنا بحسابات معقَّدة في الحالة العامّة. في مقابل ذلك نجد أنَّ الدوال المولِّدة الاحتمالية تُبدي طريقة مختصرة لتعيين توزيع مثل هذا المجموع.

المبرهنة الآتية تُمكّننا من تعيين الدالَّة المولِّدة الاحتمالية لمجموع متغيرات عشوائية مستقلّة بدلالة الدوال المولِّدة الاحتمالية لكلٍّ منها، وبدورها يمكن أن تكشف لنا عن توزيع مجموع المتغيرات العشوائية المستقلّة.

**(٣, ٢, ٣, ٧) مبرهنة**

ليكن $X_1$، $X_2$، ...و $X_n$ متغيِّرات عشوائية مستقلّة فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فعندئذ تكون العلاقة الآتية مُحقَّقة:

$$\mathbf{G}_{X_1+X_2+\dots+X_n}(s)=\prod_{k=1}^{n}\mathbf{G}_{X_k}(s)\qquad;\forall s\in\mathbb{R}\qquad[7,40]$$

البرهان: بما أنَّ المتغيِّرات العشوائية $X_1$، $X_2$،...و $X_n$ مستقلّة، فعندئذ من أجل كلّ $s \in \mathbb{R}$ ستكون المتغيِّرات العشوائية $s^{X_1}$، $s^{X_2}$،... و $s^{X_n}$ مستقلّة أيضاً، ومنه ينتج أنَّه من أجل كلّ $s \in \mathbb{R}$ سيكون الآتي مُحقَّقاً:

$$\mathbf{G}_{X_1+X_2+.....+X_n}(s) = \mathbf{E}\, s^{X_1+X_2+.....+X_n} = \mathbf{E}\left(\prod_{k=1}^{n} s^{X_k}\right) = \prod_{k=1}^{n} \mathbf{E}\, s^{X_k} = \prod_{k=1}^{n} \mathbf{G}_{X_k}(s)$$

وبهذا يتمُّ المطلوب.

### (٧,٣,٢,٤) أمثلة

١- لتكن $X_1$، $X_2$،...و $X_n$ متغيِّرات عشوائية برنوليّة مستقلّة ولكل منها المَعْلمة $p$ نفسها مع $0<p<1$، ولنأخذ المجموع الآتي من أجل $n \in \mathbb{N}$ كيفي ولكن مثَّبت:

$$S_n = X_1 + X_2 + ... + X_n \qquad (*)$$

فنجد أنَّ الدالَّة المولِّدة الاحتمالية لـ $X_k$ من أجل $k \in \mathbb{N}_n$ هي:

$$\mathbf{G}_{X_k}(s) = (1-p)\,s^0 + p\,s^1 = (1-p) + ps \qquad ;\forall s \in \mathbb{R}$$

ومن ثمَّ يكون للدالَّة المولِّدة الاحتمالية للمتغيِّر العشوائي $S_n$ العرض الآتي:

$$\mathbf{G}_{S_n}(s) = \left[\,\mathbf{G}_{X_k}(s)\,\right]^n = \left[(1-p) + ps\right]^n \qquad ;\forall s \in \mathbb{R}$$

وهذه الصيغة الأخيرة هي صيغة الدالَّة المولِّدة الاحتمالية لمتغيِّر عشوائي خاضع للتوزيع الحدَّاني بمَعْلمتين $n$ و $p$، أي إنَّ لـ $S_n$ توزيع حدَّاني $\mathbf{B}(n\,,p)$.

٢- ليكن لدينا $X_1 \sim \mathbf{B}(n_1, p)$ و $X_2 \sim \mathbf{B}(n_2, p)$ متغيِّرين عشوائيين مستقلَّين، فعندئذ نجد ما يلي:

$$\mathbf{G}_{X_1+X_2}(s) = \mathbf{G}_{X_1}(s)\cdot\mathbf{G}_{X_2}(s) = \left[(1-p) + ps\right]^{n_1+n_2} \qquad ;\forall s \in \mathbb{R}$$

وهذه الصيغة الأخيرة هي صيغة الدالَّة المولِّدة الاحتمالية لمتغيِّر عشوائي خاضع للتوزيع الحدَّاني بمَعْلمتين $n = n_1 + n_2$ و $p$، ومن ثَمَّ يكون لمجموع المتغيِّرين العشوائيين المستقلَّين المُعطيين توزيع حدَّاني بمَعْلمتين $n = n_1 + n_2$ و $p$. أي إنَّ:

$$X_1 + X_2 := X \sim \mathbf{B}(n_1 + n_2\;,\;p)$$

في الواقع تُصادفنا في بعض المسائل حالات يكون فيها عدد المتغيِّرات العشوائية $n$ نفسه هو متغيِّر عشوائي، فعندئذ الإجابة على تعيين توزيع المجموع $S_n$ في العلاقة (*) ليس بسيطاً، ولكن المبرهنة الآتية (التي سيترك برهانها كتمرين للقارئ) تقدِّم لنا إحدى الطرائق المستخدمة في حساب توزيع $S_n$ عندما تكون هذه المتغيِّرات العشوائية مستقلّة ومتطابقة في التوزيع.

### (٧,٣,٢,٥) مبرهنة

إذا كانت $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية مستقلّة فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$ ومتطابقة في التوزيع مع توزيع متغيِّر عشوائي $X$ فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$ بدالَّة مولِّدة احتمالية $\mathbf{G}_X$، وليكن $\zeta$ متغيِّراً عشوائياً بقيم في $\mathbb{N}$، ويملك دالَّة مولِّدة احتمالية $\mathbf{G}_\zeta$، فإذا كان المتغيِّر العشوائي $\zeta$ مستقل عن جميع المتغيِّرات العشوائية $X_n$ من أجل كلّ $n \in \mathbb{N}$، فعندئذ يكون للمتغيِّر العشوائي $Y = X_1 + X_2 + ... + X_\zeta$ دالَّة مولِّدة احتمالية مُعطاة من خلال العلاقة الآتية:

$$\mathbf{G}_Y(s) = \mathbf{G}_\zeta\left(\mathbf{G}_X(s)\right) \qquad ;\forall s \in \mathbb{R}$$ [7,41]

### (٧,٣,٢,٦) ملاحظة

إنَّ لهذه العلاقة الأخيرة أهميّة كبيرة في التطبيقات، ولكن من الممكن أن تصبح مُربكةً (صعبة الحل) عندما تكون فيها الدوال $\mathbf{G}_\zeta$

و $\mathbf{G}_X$ مُركَّبة بشكل مُعقَّد، وسنتجنبها بوضع $P(\zeta = n) = 1$ ، فعندئذ سيكون من أجل كلّ $s \in \mathbb{R}$ لدينا ما يلي مُحقَّقاً (سنقبل هذه النتائج دون برهان):

**a)** $\mathbf{G}_\zeta(s) = s^n$ [7,42-a]

**b)** $\mathbf{G}_{S_n}(s) = (\mathbf{G}_X(s))^n$ [7,42-b]

والمثال الآتي يوضِّح لنا مدى أهميّة المبرهنة السابقة في التطبيقات.

**مثال (٧,٣,٢,٧)**

يرد إلى مركز صيانة للسّيارات $N$ سيارة في اليوم، علماً أنَّ $N$ يخضع لتوزيع بواسون بمَعْلمة $\lambda > 0$، ولنفترض أنَّ كلّ سيارة يمكن إتمام صيانتها خلال اليوم نفسه باحتمال $0 < p < 1$، وأنَّ عمليّة الصيانة للسيارات تتمُّ بشكل مستقل بعضها عن بعضها الآخر، ولنفترض أيضاً أنَّ $X_k$ مع $k \in \mathbb{N}_N$ هو متغيِّر عشوائي راصد للسيارة ذات الرقم $k$ التي دخلت المركز في اليوم نفسه، فنجد أنَّ $X_1$، $X_2$،...و $X_N$ هي متغيِّرات عشوائية برنولية مستقلّة بمَعْلمة $p$ ، ومن ثَمَّ المتغيِّر العشوائي $Y = X_1 + X_2 + ... + X_N$ سيكون دالاً على عدد السيارات التي تمَّ إصلاحها في اليوم نفسه، وهنا نلاحظ أنَّه لم يَعد بإمكاننا أن ندعي أنَّ لـ $Y$ توزيع حدَّاني $\mathbf{B}(N, p)$ لأنَّ $N$ أصبح متغيِّراً عشوائياً. لذلك سنستخدم الدراسة الأخيرة في تعيين توزيع المتغيِّر العشوائي $Y$.

**الحـل:** من مُعطيات المسألة لدينا الآتي:

$$\mathbf{G}_N(s) = \sum_{k=0}^{\infty} s^k \frac{\lambda^k}{k!} \mathbf{e}^{-\lambda} = \mathbf{exp}[\lambda(s-1)] \qquad ;\forall s \in \mathbb{R}$$

ومن ثَمَّ يكون من أجل كلّ $k \in \mathbb{N}_N$ و $s \in \mathbb{R}$ لدينا ما يلي:

$$\mathbf{G}_{X_k}(s) = (1-p) + ps = sp + (1-p) = p(s-1) + 1$$

وهكذا يصبح من أجل كلّ $k \in \mathbb{N}_N$ و $s \in \mathbb{R}$ لدينا الآتي مُحقّقاً:

$$\mathbf{G}_Y(s) = \mathbf{G}_N(\mathbf{G}_{X_k}(s)) = \mathbf{exp}\{\lambda p \cdot (s-1)\}$$

وبمقارنتها مع $\mathbf{G}_N$ نجد أنَّ $\mathbf{G}_Y$ هي الدالَّة المولِّدة الاحتمالية لمتغيِّر عشوائي بواسوني بمَعْلمة $\lambda \cdot p$ ، ومن ثَمَّ سيكون للمتغيِّر العشوائي $Y$ توزيع بواسوني بمَعْلمة $\bar{\lambda} = \lambda \cdot p$ .

كتطبيق على هذا المثال سنفترض تحت الفرضيات الواردة آنفًا أنَّه يدخل إلى هذا المركز 100 سيارة في المتوسط، وأنَّ احتمال إصلاح سيارة في اليوم نفسه يساوي 0.75 ، ولنحسب احتمال إصلاح 80 سيارة من السيارات التي دخلت هذا المركز في اليوم نفسه.

لدينا من مُعطيات المسألة $p = 0.75$ و $\lambda = \mathbf{E}N = 100$، ولكن من أجل مسألتنا هذه لدينا علاوةً على ما سبق ما يلي:

$$P(S_N = k) = \frac{(\lambda p)^k}{k!} \mathbf{e}^{-\lambda p}$$

ومنه يكون الاحتمال المطلوب هو:

$$P(S_N = 80) = \frac{(75)^{80}}{(80)!} \mathbf{e}^{-75} = 0.0379$$

أي إنَّ احتمال إصلاح 80 سيارة من أصل 100 سيارة دخلت المركز في اليوم نفسه يساوي 0.0379.

### (٧,٣,٣) الدالَّة المولِّدة لعزوم متغيِّر عشوائي Moment Generating Function of a Random Variable

إنَّ الدوال المولِّدة الاحتمالية مفيدة في دراسة المتغيِّرات العشوائية الصحيحة غير السالبة، ولكن من أجل متغيِّرات عشوائية أكثر عموميَّة يُفضَّل استخدام التحويل $s = \mathbf{e}^t$ مع $t \in \mathbb{R}$ في صيغة الدالَّة المولِّدة للعزوم $\mathbf{G}_X$ فنحصل بذلك على ما يُسمَّى بالدالَّة المولِّدة للعزوم التي تمكننا في كثير من الحالات من حساب العزوم الابتدائية للمتغيِّرات العشوائية من خلال اشتقاق دوالها، وهذا المفهوم الجديد يُقدِّمه لنا التعريف الآتي.

#### (٧,٣,٣,١) تعريف (الدالَّة المولِّدة لعزوم متغيِّر عشوائي)

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، ولنعرِّف من أجل هذا المتغيِّر العشوائي دالَّة حقيقية $\mathbf{M}_X$ على $\mathbb{R}$ من خلال العلاقة الآتية:

$$\mathbf{M}_X : \mathbb{R} \longrightarrow \mathbb{R} \;\; ; t \mapsto \mathbf{M}_X(t) = \mathbf{E}\, \mathbf{e}^{tX} \qquad [7,43\text{-a}]$$

فعندئذ تُدعى $\mathbf{M}_X$ الدالَّة المولِّدة لعزوم المتغيِّر العشوائي $X$.

#### (٧,٣,٣,٢) ملاحظات

١- فإذا كان $X$ متغيِّراً عشوائياً متقطِّعاً مجموعة قيمه $\mathbf{X} = \{x_i \;\; ; i \in I\}$، فعندئذ يكون للدالَّة المولِّدة لعزوم $X$ العرض الآتي:

$$\mathbf{M}_X(t) = \sum_{i \in I} \mathbf{e}^{t x_i} P(X = x_i) \qquad ; t \in \mathbb{R} \qquad [7,43\text{-b}]$$

وفي حال كانت مجموعة قيمه $\mathbf{X} = \mathbb{N}$، فعندئذ سيصبح للدالَّة المولِّدة لعزوم $X$ العرض الآتي:

$$\mathbf{M}_X(t) = \sum_{k \in \mathbb{N}} \mathbf{e}^{tk} P(X = k) \qquad ; t \in \mathbb{R} \qquad [7,43\text{-c}]$$

الآن، لكي يملك متغيِّرا عشوائيًّا $X$ دالَّة مولِّدة للعزوم (أي إنَّ $\mathbf{M}_X$ موجودة) يجب أن تكون أيّ من المتسلسلتين العدديتين السابقتين [7,43-b] و [7,43-c] متقاربةً مطلقاً.

٢- إمّا إذا كان $X$ متغيِّراً عشوائياً مستمراً بدالَّة كثافة احتمالية $f_X$، فعندئذ يكون للدالَّة المولِّدة لعزوم $X$ العرض الآتي:

$$\mathbf{M}_X(t) = \int_{-\infty}^{\infty} \mathbf{e}^{tx} f_X(x)\, dx \qquad ; t \in \mathbb{R} \qquad [7,43\text{-d}]$$

ولكي يملك متغيِّر عشوائي $X$ دالَّة مولِّدة للعزوم يجب أن يكون التكامل المعتّل في العلاقة [7,43-d] متقارباً مطلقاً.

٣- نلاحظ الصلة الوطيدة بين $\mathbf{M}_X$ و تحويل لابلاس Laplace Transform، وذلك لأنَّه لدينا:

$$\mathbf{M}_X(t) = \int_{-\infty}^{+\infty} \mathbf{e}^{tX}\, dF_X(x) \qquad ; t \in \mathbb{R} \qquad [7,44]$$

وإذا كان $X$ مستمراً بدالَّة كثافة احتمالية $f_X$ فعندئذ يصبح للعلاقة السابقة العرض الآتي:

$$\mathbf{M}_X(t) = \int_{-\infty}^{+\infty} \mathbf{e}^{tX} f_X(x)\, dx \qquad ; t \in \mathbb{R}$$

٤- إنَّ الدوال المولِّدة للعزوم تمتلك خصائص مماثلة لخصائص الدوال المولِّدة الاحتمالية (لأنَّها حالة خاصّة من الدوال المولِّدة الاحتمالية عندما نأخذ $s = \mathbf{e}^t$)، فإذا كانت $\mathbf{M}_X$ موجودة على فترات مفتوحة تحوي نقطة الأصل (المبدأ)، فعندئذ:

أ- سيكون لدينا:

$$\mathbf{E}X^{k}=\mathbf{M}_X^{(k)}(0) \qquad ;k\in\mathbb{N} \tag{7,45}$$

ب- سيكون لدينا من نشر تايلور Taylor's Expansion ما يلي:

$$\mathbf{M}_X(t)=\sum_{k=0}^{\infty}\frac{\mathbf{E}X^{k}}{k\,!}t^{k} \tag{7,46}$$

ج- إذا كان $Y$ متغيِّراً عشوائياً فوق الفضاء الاحتمالي $[\Omega,\mathcal{A},P]$ ومستقلاً عن $X$، فعندئذ ينتج (إثباتها سيترك كتمرين) لدينا:

$$\mathbf{M}_{X+Y}(t)=\mathbf{M}_X(t)\cdot\mathbf{M}_Y(t) \qquad ;t\in\mathbb{R} \tag{7,47}$$

وهذه النتيجة يمكن تعميمها على أي عدد منته من المتغيِّرات العشوائية المستقلّة، فلو كانت لدينا $X_1$، $X_2$، ...و $X_n$ متغيِّرات عشوائية مستقلّة فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$، فعندئذ ستكون العلاقة الآتية محقَّقةً:

$$\mathbf{M}_{X_1+X_2+\dots+X_n}(t)=\prod_{k=1}^{n}\mathbf{M}_{X_k}(t) \qquad ;t\in\mathbb{R} \tag{7,48}$$

### (٧,٣,٣,٣) أمثلة

١- لنأخذ $X\sim\mathbf{D}(c)$ متغيِّراً عشوائياً، فعندئذ يكون للدالَّة المولِّدة لعزوم $X$ العرض الآتي:

$$\mathbf{M}_X(t)=\mathbf{e}^{tc} \qquad ;\forall t\in\mathbb{R}$$

وذلك لأنَّه من أجل كلّ $t\in\mathbb{R}$ لدينا:

$$\mathbf{M}_X(t)=\mathbf{E}_X\left(\mathbf{e}^{tX}\right)=\sum_{i\in I}\mathbf{e}^{tx_i}P(X=x_i)=\mathbf{e}^{tc}\cdot P(X=c)=\mathbf{e}^{tc}$$

فنلاحظ أنَّ الدالَّة المولِّدة لعزوم هذا المتغيِّر العشوائي قابلة للاشتقاق على $\mathbb{R}$ عدداً كيفياً من المرَّات، ومن ثمَّ يمكننا حساب العزوم الابتدائية من مختلف المراتب لهذا المتغيِّر العشوائي حيث لدينا من العلاقة [7,45] ما يلي:

$$\mathbf{E}X=\mathbf{M}_X^{(1)}(t)\Big|_{t=0}=\left(\mathbf{e}^{tc}\right)'\Big|_{t=0}=c\,\mathbf{e}^{tc}\Big|_{t=0}=c$$

وبالمثل نجد أنَّ العزوم الابتدائي من المرتبة الثانية لهذا المتغيِّر العشوائي يساوي:

$$\mathbf{E}X^{2}=\mathbf{M}_X^{(2)}(t)\Big|_{t=0}=\left(\mathbf{e}^{tc}\right)''\Big|_{t=0}=c^{2}\mathbf{e}^{tc}\Big|_{t=0}=c^{2}$$

٢- لنأخذ $X\sim\mathbf{B}(n,p)$ متغيِّراً عشوائياً، فعندئذ يكون للدالَّة المولِّدة لعزوم $X$ العرض الآتي:

$$\mathbf{M}_X(t)=\left[p(\mathbf{e}^{t}-1)+1\right]^{n} \qquad ;\forall t\in\mathbb{R}$$

وذلك لأنَّه من أجل كلّ $t\in\mathbb{R}$ لدينا:

$$\mathbf{M}_X(t)=\mathbf{E}_X\left(\mathbf{e}^{tX}\right)=\sum_{i\in I}\mathbf{e}^{tx_i}P(X=x_i)=\sum_{k=0}^{n}\mathbf{e}^{tk}P(X=k)=\sum_{k=0}^{n}\mathbf{e}^{tk}\cdot\binom{n}{k}\cdot p^{k}\cdot(1-p)^{n-k}$$
$$=\sum_{k=0}^{n}\binom{n}{k}\cdot(\mathbf{e}^{t}p)^{k}\cdot(1-p)^{n-k}=\left[\mathbf{e}^{t}p+(1-p)\right]^{n}=\left[p(\mathbf{e}^{t}-1)+1\right]^{n}$$

فنلاحظ أنَّ الدالَّة المولِّدة لعزوم هذا المتغيِّر العشوائي قابلة للاشتقاق على $\mathbb{R}$ عدداً كيفياً من المرَّات، ومن ثمَّ يمكننا حساب العزوم

الابتدائية من مختلف المراتب لهذا المتغيِّر العشوائي حيث لدينا من العلاقة [7,45] ما يلي:

$$\mathbf{E}X=\mathbf{M}_X^{(1)}(t)\Big|_{t=0}=\left(\left[p(\mathbf{e}^t-1)+1\right]^n\right)'\Bigg|_{t=0}=n\,p\left[p(\mathbf{e}^t-1)+1\right]^{n-1}\Bigg|_{t=0}=n\,p$$

وبالمثل نجد أنَّ العزوم الابتدائي من المرتبة الثانية لهذا المتغيِّر العشوائي يساوي:

$$\mathbf{E}X^2=\mathbf{M}_X^{(2)}(t)\Big|_{t=0}=\left(\left[p(\mathbf{e}^t-1)+1\right]^n\right)''\Bigg|_{t=0}=\left(n\,p\left[p(\mathbf{e}^t-1)+1\right]^{n-1}\right)'\Bigg|_{t=0}$$

$$=n(n-1)p^2\left[p(\mathbf{e}^t-1)+1\right]^{n-1}\Bigg|_{t=0}=n(n-1)p^2$$

ومن ثَمَّ نجد باستخدام صيغة شتاينر أنَّ التباين لهذا المتغيِّر العشوائي يساوي:

$$\mathbf{var}\,X=\mathbf{E}X^2-(\mathbf{E}X)^2=n(n-1)p^2-(n\,p)^2=n\,p\,(1-p)$$

لاحظ هنا السرعة والسهولة في حساب العزوم الابتدائي من المرتبة الأولى والثانية لهذا المتغيِّر العشوائي مقارنة بما سبق وفقاً لاستخدام تعريف العزوم الابتدائية.

٣– لنأخذ $X\sim\mathbf{P}(\lambda)$ متغيِّراً عشوائياً، فعندئذ يكون للدالَّة المولِّدة لعزوم $X$ العرض الآتي:

$$\mathbf{M}_X(t)=\mathbf{exp}\left\{\lambda(\mathbf{e}^t-1)\right\}\qquad;\forall t\in\mathbb{R}$$

وذلك لأنَّه من أجل كلّ $t\ni\mathbb{R}$ لدينا:

$$\mathbf{M}_X(t)=\mathbf{E}\,\mathbf{e}^{tX}=\sum_{k=0}^{\infty}\mathbf{e}^{tk}P(X=k)=\sum_{k=0}^{\infty}\mathbf{e}^{tk}\frac{\lambda^k}{k!}\mathbf{e}^{-\lambda}$$

$$=\mathbf{e}^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda)^k}{k!}=\mathbf{e}^{-\lambda}\mathbf{e}^{\lambda\mathbf{e}^t}=\mathbf{exp}\left\{\lambda(\mathbf{e}^t-1)\right\}$$

وأمّا استنتاج العزوم الابتدائية من المرتبة الأولى والثانية لهذا المتغيِّر العشوائي فيترك للقارئ.

٤– لنأخذ $X\sim\mathbf{U}([a,b])$ متغيِّراً عشوائياً، فعندئذ يكون للدالَّة المولِّدة لعزوم $X$ العرض الآتي:

$$\mathbf{M}_X(t)=\frac{\mathbf{e}^{tb}-\mathbf{e}^{ta}}{t(b-a)}\qquad;\forall t\in\mathbb{R}$$

وذلك لأنَّه من أجل كلّ $t\ni\mathbb{R}$ لدينا:

$$\mathbf{M}_X(t)=\int_{-\infty}^{\infty}\mathbf{e}^{tx}f_X(x)\,dx=\frac{1}{b-a}\int_a^b\mathbf{e}^{tx}\,dx=\frac{1}{(b-a)}\frac{\mathbf{e}^{tx}}{\mathbf{ln}\,\mathbf{e}^t}\Bigg|_a^b=\frac{\mathbf{e}^{tb}-\mathbf{e}^{ta}}{t(b-a)}$$

وأمّا استنتاج العزوم الابتدائية من المرتبة الأولى والثانية لهذا المتغيِّر العشوائي فيترك للقارئ.

٥– لنأخذ $X\sim\mathbf{EX}(\lambda)$ متغيِّراً عشوائياً، فعندئذ يكون للدالَّة المولِّدة لعزوم $X$ العرض الآتي:

$$\mathbf{M}_X(t)=\frac{\lambda}{\lambda-t}\qquad;\forall t\in\mathbb{R}$$

وذلك لأنَّه من أجل كلّ $t\ni\mathbb{R}$ لدينا:

$$\mathbf{M}_X(t) = \int_{-\infty}^{\infty} \mathbf{e}^{tx} f_X(x)\, dx = \int_{-\infty}^{\infty} \mathbf{e}^{tx} \lambda \mathbf{e}^{-\lambda x} I_{[0,\infty)}(x)\, dx = \lambda \int_{0}^{\infty} \mathbf{e}^{tx} \mathbf{e}^{-\lambda x}\, dx$$

$$= \lambda \int_{0}^{\infty} \mathbf{e}^{-x(\lambda - t)}\, dx = \frac{\lambda}{-(\lambda - t)} \mathbf{e}^{-x(\lambda - t)} \Big|_0^{\infty} = \frac{\lambda}{\lambda - t}$$

وأمّا استنتاج العزوم الابتدائية من المرتبة الأولى والثانية لهذا المتغيِّر العشوائي فيترك للقارئ.

## (٧,٤) الدوال المميزة للتوزيعات الاحتمالية

**Characteristic Function of Probability Distributions**

لقد أثبتت الدراسات أنَّ الدوال المميِّزة لدوال التوزيع الاحتمالية لها شأن كبير جداً في مجال الدراسات التحليلية لنظرية الاحتمالات، لا بل أكثر من ذلك، يمكن اشتقاق الكثير من المبرهنات المهمّة في مجال الطرائق التحليلية لنظرية الاحتمالات بوساطة الدوال المميِّزة، ولذلك لُجِئ إلى استخدام هذا النوع من الدوال في حل العديد من القضايا المتعلّقة بهذه النظرية، وعلى وجه الخصوص تلك القضايا التي تهتّم بتوليد العزوم للمتغيِّرات العشوائية. علاوةً على ما سبق، من النتائج المهمّة والرئيسية للدراسات التحليلية لنظرية الاحتمالات الموضوعات الآتية:

- نظريات الوحدانية Uniqueness theorem's.
- نظريات العكس Inversion theorem's.
- نظريات التلاف Convolution theorem's.
- نظريات الاستمرار Continuity theorem's.
- نظريات النهاية المركزية Central limit theorem's.

إنَّ هذه الدراسات تسمح لنا باستبدال الدراسة على دالّة توزيع احتمالية باستقصاء مكافئ لدوال مميِّزة مقابلة لها، وكذلك نجد أنَّ الدوال المميِّزة للمتغيِّرات العشوائية تُمكّننا من تعيين التوزيعات التي لا تملك عزوماً ابتدائية أيضاً.

### (٧, ٤, ١) الدالّة المميِّزة لتوزيع احتمالي Characteristic Function of Probability Distribution

سنقوم فيما يلي بتقديم تعريف المتغيِّر العشوائي العقدي أولاً، ومن ثمَّ تعريف الدالّة المميِّزة لدالّة توزيع احتمالية، ونستخلص منها الدالّة المميِّزة لمتغيِّر عشوائي، علماً أنَّ دراستنا التالية ستكون مركَّزة على الدوال المميِّزة للمتغيِّرات العشوائية.

#### (٧, ٤, ١, ١) تعريف (المتغيِّر العشوائي العقدي Complex Random Variable)

ليكن $X$ و$Y$ متغيِّرين عشوائيين فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، فعندئذ يُدعى المتغيِّر العشوائي:

$$Z := X + \mathbf{i}\, Y \qquad [7,49]$$

**متغيِّراً عشوائياً عقدياً (أو مركَّباً)**. علماً أنَّ $\mathbf{i}$ هو العدد العقدي الممثَّل بالنقطة $(0,1)$ في المستوي العقدي، والمميَّز بالعلاقة $\mathbf{i}^2 = -1$.

من الواضح هنا أنَّه يكون لهذا المتغيِّر العشوائي توقع رياضياتي **إذا وفقط إذا** كان $\mathbf{E}|X| < +\infty$ وكذلك $\mathbf{E}|Y| < +\infty$، وفي هذه الحالة يكون لدينا:

$$\mathbf{E}Z = \mathbf{E}X + \mathbf{i} \cdot \mathbf{E}Y \qquad [7,50]$$

إنَّ الخصائص المتعلّقة بالتوقع الرياضياتي للمتغيِّرات العشوائية الحقيقية تبقى سارية المفعول من أجل المتغيِّرات العشوائية العقدية مع الانتباه لدى التعامل مع المتباينات، وذلك لأنَّه يجب أخذ طويلة المتغيِّر العشوائي العقدي في هذه الحالة.

**(٢ , ١ , ٤ , ٧) تعريف (الدالَّة المميِّزة لدالَّة توزيع احتمالية) Characteristic Function of Distribution Function**

لتكن $F$ دالة توزيع احتمالية، ولنعرِّف من أجلها دالَّة عقدية $\varphi_F$ على $\mathbb{R}$ من خلال العلاقة الآتية:

$$\varphi_F : \mathbb{R} \longrightarrow \mathbb{C} \quad ; t \mapsto \varphi_F(t) := \int_{-\infty}^{\infty} \mathbf{e}^{\mathbf{i}tx} \, dF(x) \qquad [7,51]$$

عندئذ تُدعى $\varphi_F$ **الدالَّة المميِّزة** للدالَّة $F$.

**(٣ , ١ , ٤ , ٧) ملاحظات**

١- نلاحظ هنا أنَّ التكامل المقدَّم في العلاقة [7,51] هو من نوع تكاملات ريمان- ستلجس المعتلَّة، وهذا التكامل موجود من أجل كلّ دالَّة توزيع احتمالية $F$ لأنَّ $\left| \mathbf{e}^{\mathbf{i}tx} \right| = 1$ .

٢- إذا كانت $F_X$ دالَّة توزيع متغيِّر عشوائي $X$ فعندئذ سنكتب $\varphi_X$ عوضاً عن $\varphi_{F_X}$ ، ونجد أنَّ التكامل الذي في الطرف الأيمن للعلاقة [7,51] هو التوقع الرياضياتي للمتغيِّر العشوائي العقدي $\mathbf{e}^{\mathbf{i}tX}$ ، ومن ثمَّ يمكننا أن نكتب الآتي:

$$\varphi_X(t) = \mathbf{E}\, \mathbf{e}^{\mathbf{i}tX} \qquad ; \forall t \in \mathbb{R} \qquad [7,52]$$

وهذه الدالَّة $\varphi_X$ تُدعى **الدالَّة المميِّزة للمتغيِّر العشوائي** $X$ .

٣- إنَّ الدوال المميِّزة للمتغيِّرات العشوائية ذات صلة وطيدة بـ تحويل فورييه Fourier Translation، وذلك لأنَّ التكامل الذي في الطرف الأيمن من العلاقة [7,51] يُمثِّل تحويل فورييه – ستلجلس Fourier- Stieltjes Translation ، وأمّا التكامل:

$$F(t) = \int_{-\infty}^{\infty} \mathbf{e}^{-\mathbf{i}tx} f(x)\, dx$$

فإنَّه يمثِّل تحويل فورييه التكاملي للدالَّة $f(x)$، في حين أنَّ التكامل:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathbf{e}^{\mathbf{i}tx} F(t)\, dt$$

هو تحويل فورييه العكسي.

٤- من أجل كلّ $t \in \mathbb{R}$ يمكن عرض الدالَّة $\varphi_F$ كما في العلاقة الآتية:

$$\varphi_F(t) = \int_{-\infty}^{\infty} \cos tx \; dF(x) + \mathbf{i} \int_{-\infty}^{\infty} \sin tx \; dF(x) \qquad ; \forall t \in \mathbb{R} \qquad [7,53]$$

ومن العلاقة الأخيرة [7,53] نستنتج أنَّ الدالَّة المميِّزة $\varphi_F$ موجودةً في جوار المبدأ من أجل كلّ دالة توزيع احتمالية $F$، وإذا كانت الدالَّة $F$ هي دالَّة توزيع متغيِّر عشوائي $X$ ، فعندئذ سيكون لدينا:

$$\varphi_X(t) = \mathbf{E}(\cos tX) + \mathbf{i}\,\mathbf{E}(\sin tX) \qquad [7,54]$$

**(٢ , ٤ , ٧) خصائص الدالَّة المميِّزة لتوزيع احتمالي**

تقدِّم لنا المبرهنة الآتية بعض الخصائص المهمّة للدوال المميِّزة للمتغيِّرات العشوائية.

**(١ , ٢ , ٤ , ٧) مبرهنة**

لتكن $\varphi_X$ دالَّة مميِّزة لمتغيِّر عشوائي $X$ فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ بدالَّة توزيع احتمالية $F_X$ ، فعندئذ ينتج لدينا من التعريف أنَّ $\varphi_X(0) = \mathbf{E}(1) = 1$ ، وكذلك من أجل كلّ $t \in \mathbb{R}$ يكون لهذه الدالَّة الخصائص الآتية:

١ - لدينا:

$$\text{a)}\ \left|\varphi_X(t)\right|\le 1 \qquad \& \qquad \text{b)}\ \varphi_X(-t)=\overline{\varphi_X(t)}$$

٢ - الدالّة $\varphi_X$ موجبة التعيين Positive Definite، والتي تعني أنّه من أجل كلّ $i \in \mathbb{N}_n$ وأيّة قيم $z_i \in \mathbb{C}$ و $t_i \in \mathbb{R}$ لدينا العلاقة الآتية مُحقَّقة:

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\varphi_X(t_j-t_k)\,z_j\,\bar{z}_k \ge 0 \qquad [7,55]$$

**البرهان:** من أجل الفقرة (١) نلاحظ وضوحاً أنّه من أجل كلّ $t \in \mathbb{R}$ لدينا الآتي محقّقاً:

$$\left|\varphi_X(t)\right| \le \int_{\mathbb{R}}\left|\mathbf{e}^{\mathbf{i}tx}\right| dF_X(t)=\int_{\mathbb{R}} dF_X(t)=1$$

وأخيراً من أجل كلّ $t \in \mathbb{R}$ لدينا:

$$\varphi_X(-t) = \int_{\mathbb{R}}\mathbf{e}^{-\mathbf{i}tx}\ dF_X(t) = \int_{-\infty}^{\infty}\mathbf{cos}\,tx\ dF_X(x)-\mathbf{i}\int_{-\infty}^{\infty}\mathbf{sin}\,tx\ dF_X(x)=\overline{\varphi_X(t)}$$

٢ - من أجل الفقرة (٢) نجد من أجل كلّ $i \in \mathbb{N}_n$ وأيّة قيم $z_i \in \mathbb{C}$ و $t_i \in \mathbb{R}$ ما يلي مُحقَّقاً:

$$\sum_{j=1}^{n}\sum_{k=1}^{n}\varphi_X(t_j-t_k)\,z_j\,\bar{z}_k=\sum_{j=1}^{n}\sum_{k=1}^{n}\int_{\mathbb{R}} z_j\,\mathbf{e}^{\mathbf{i}t_jx}\,\bar{z}_k\,\mathbf{e}^{-\mathbf{i}t_kx}\ dF_X(x)=\mathbf{E}\left|\sum_{j\in\mathbb{N}_n} z_j\ \mathbf{exp}\,(\mathbf{i}\,t_j\,x)\right|^2 \ge 0$$

وبهذا يتمّ البرهان.

تظهر لنا المبرهنة الآتية (التي سنقدّمها دون برهان لأنّها تُمثّل مبرهنة تايلور من أجل الدوال العقدية) بعض الخصائص الإضافية للدوال المميّزة للمتغيّرات العشوائية.

### (٢,٢,٤,٧) مبرهنة

ليكن $X$ متغيّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$، فعندئذ:

١ - إذا كان المقدار $\varphi_X^{(k)}(0)$ موجوداً فإنّه:

أ- إذا كان $k$ عدداً زوجياً، فإنّه سيكون لدينا $\mathbf{E}\left|X^k\right|<+\infty$.

ب- إذا كان $k$ عدداً فردياً، فإنّه سيكون لدينا $\mathbf{E}\left|X^{k-1}\right|<+\infty$.

٢- إذا كان $\mathbf{E}\left|X^k\right|<+\infty$، فعندئذ من أجل كلّ $t \in \mathbb{R}$ يكون لدينا:

$$\varphi_X(t)=\sum_{j=0}^{k}\frac{\mathbf{E}\,X^j}{j\,!}(\mathbf{i}\,t)^j+\theta(t^k) \qquad [7,56]$$

علماً أنّ $\lim\limits_{t\to 0}\theta(t^k)=0$.

### (٣,٢,٤,٧) نتيجة

من المبرهنة السابقة نستنتج تحت الفرض أنّه من أجل $k \in \mathbb{N}$ لدينا $\mathbf{E}\left|X^k\right|<+\infty$، فعندئذ سينتج لدينا أنّ:

$$\varphi_X^{(k)}(0)=\mathbf{i}^k\cdot\mathbf{E}\,X^k \qquad [7,57]$$

وهذه العلاقة الأخيرة تساعدنا في حساب لنا العزوم الابتدائية حتى المرتبة $k$ للمتغيّر العشوائي $X$ بوساطة الاشتقاق للدالّة $\varphi_X$.

قبل ختام فقرة الدوال المميِّزة للمتغيِّرات العشوائية سنقدِّم مبرهنتين تعرضان لنا بعض الخصائص الإضافية التي تتمتّع بها هذه الدوال.

**(٧, ٤, ٢, ٤) مبرهنة**

ليكن $X$ و $Y$ متغيِّرين عشوائيين فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فعندئذ:

١- إذا كان $Z = aX + b$ مع $a, b\in\mathbb{R}$ و $a \neq 0$ فإنَّه سيكون لدينا:

$$\varphi_Z(t) = \mathbf{e}^{\mathbf{i}tb}\cdot\varphi_X(at) \quad ;\forall t\in\mathbb{R} \qquad [7,58]$$

٢- إذا كان $Y$ مستقلاً عن $X$، فعندئذ يكون لدينا:

$$\varphi_{X+Y}(t) = \varphi_X(t)\cdot\varphi_Y(t) \quad ;\forall t\in\mathbb{R} \qquad [7,59\text{-a}]$$

**البرهان:** لإثبات صحة الفقرة (١) لدينا من أجل $t\in\mathbb{R}$ ما يلي:

$$\varphi_Z(t) = \mathbf{E}\ \mathbf{e}^{\mathbf{i}t(aX+b)} = \mathbf{E}\left(\mathbf{e}^{\mathbf{i}tb}\cdot\mathbf{e}^{\mathbf{i}atX}\right) = \mathbf{e}^{\mathbf{i}tb}\cdot\mathbf{E}\ \mathbf{e}^{\mathbf{i}(at)X} = \mathbf{e}^{\mathbf{i}tb}\cdot\varphi_X(at)$$

وأمّا لإثبات صحة الفقرة (٢) فإنَّه بسبب استقلال $Y$ عن $X$ سيكون من أجل كل $t\in\mathbb{R}$ لدينا ما يلي:

$$\varphi_{X+Y}(t) = \mathbf{E}\ \mathbf{e}^{\mathbf{i}t(X+Y)} = \mathbf{E}\left(\mathbf{e}^{\mathbf{i}tX}\cdot\mathbf{e}^{\mathbf{i}tY}\right) \overset{\text{من استقلال } X \text{ و} Y}{=\!=\!=\!=\!=} \mathbf{E}\ \mathbf{e}^{\mathbf{i}tX}\cdot\mathbf{E}\ \mathbf{e}^{\mathbf{i}tY} = \varphi_X(t)\cdot\varphi_Y(t)$$

وبهذا يتمّ المطلوب.

**(٧, ٤, ٢, ٥) ملاحظات**

١- إنَّ البند الثاني من المبرهنة السابقة يمكن تعميمها على أي عدد منته من المتغيِّرات العشـــوائية المستقلّة (إثباتها سيترك كتمرين للقارئ)، أي إنَّه إذا كانت $X_1$، $X_2$، ...و $X_n$ متغيِّرات عشوائية مستقلّة فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، فعندئذ من أجل كل $t\in\mathbb{R}$ ستكون العلاقة الآتية مُحَقَّقة:

$$\varphi_{X_1+X_2+\ldots+X_n}(t) = \prod_{k=1}^{n}\varphi_{X_k}(t) \qquad [7,59\text{-b}]$$

٢- يجب الانتباه لدى حساب التكاملات في مسائل تتعلّق بالدوال المميِّزة؛ وذلك لأنّنا نتعامل مع تكاملات عقدية (ليست حقيقية)، ومن ثم يجب الانتباه لوضع النقطة $(0,1)$ الممثلة للعدد العقدي $\mathbf{i}$ بالنسبة للقطاع المُكامل عليه.

**(٧, ٤, ٢, ٦) أمثلة**

١- ليكن $X\sim\mathbf{B}(p)$، فعندئذ نجد من أجل كلّ $t\in\mathbb{R}$ ما يلي:

$$\varphi_X(t) = \mathbf{E}\ \mathbf{e}^{\mathbf{i}tX} = \sum_{k=0}^{1}\mathbf{e}^{\mathbf{i}tk}P(X=k) = (1-p) + p\ \mathbf{e}^{\mathbf{i}t}$$

٢- ليكن $X\sim\mathbf{B}(n,p)$، فعندئذ، وكما أوضحنا سابقاً، يمكن أخذ توزيع هذا المتغيِّر العشوائي كتوزيع مجموع $n$ متغيِّر عشوائي برنولي مستقلّ $Y_1$، $Y_2$، ...و $Y_n$، ومن ثمَّ يكون من أجل كلّ $t\in\mathbb{R}$ لدينا ما يلي مُحَقَّقاً:

$$\varphi_X(t) = \varphi_{Y_1+Y_2+\ldots+Y_n}(t) = \varphi_{Y_1}(t)\cdot\varphi_{Y_2}(t)\cdot\ldots\cdot\varphi_{Y_n}(t) = \left[(1-p)+p\ \mathbf{e}^{\mathbf{i}t}\right]^n$$

٣- ليكن $X \sim \mathbf{EX}(\lambda)$، فعندئذ سيكون لدينا $\varphi_X(t) = \frac{\lambda}{\lambda - \mathbf{i}\, t}$ من أجل كل $t \ni \mathbb{R}$، وذلك لأنَّه من أجل كلّ $t \ni \mathbb{R}$ لدينا ما يلي:

$$\varphi_X(t) = \mathbf{E}\mathbf{e}^{\mathbf{i}tX} = \int_{-\infty}^{\infty} \lambda\, \mathbf{e}^{\mathbf{i}tx}\, \mathbf{e}^{-\lambda x}\, \mathbf{I}_{[0,\infty)}(x)\, dx = \lambda \int_{0}^{\infty} \mathbf{e}^{-\lambda x} (\mathbf{cos}\, t\, x + \mathbf{i}\cdot\mathbf{sin}\, t\, x)\, dx$$
$$= \lambda \left[ \int_{0}^{\infty} \mathbf{e}^{-\lambda x}\, \mathbf{cos}\, t\, x\, dx + \mathbf{i} \int_{0}^{\infty} \mathbf{e}^{-\lambda x}\, \mathbf{sin}\, t\, x\, dx \right]$$

والتي نجد منها أنَّ:

$$\varphi_X(t) = \mathbf{E}\mathbf{e}^{\mathbf{i}tX} = \frac{\lambda^2}{\lambda^2 + t^2} + \mathbf{i} \frac{\lambda t}{\lambda^2 + t^2} = \frac{\lambda}{\lambda - \mathbf{i}t}$$

٤- ليكن $X \sim \mathbf{N}(\mu,\sigma)$، فعندئذ سيكون لدينا $\varphi_X(t) = \mathbf{e}^{\mathbf{i}\mu t - \frac{1}{2}\sigma^2 t^2}$ من أجل كل $t \ni \mathbb{R}$، ولإثبات ذلك سنأخذ $Y$ متغيِّراً عشوائياً خاضعاً للتوزيع الطبيعي المعياري $Y \sim \mathbf{N}(0,1)$، فعندئذ من أجل كلّ $t \ni \mathbb{R}$ يكون لدينا ما يلي:

$$\varphi_Y(t) = \mathbf{E}\, \mathbf{e}^{\mathbf{i}tY} = \int_{-\infty}^{\infty} \mathbf{e}^{\mathbf{i}t\,y} \cdot \frac{1}{\sqrt{2\pi}}\, \mathbf{e}^{-\frac{y^2}{2}}\, dy = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, \mathbf{exp}\left( \mathbf{i}t\,y - \frac{y^2}{2} \right) dy$$

وبحساب هذا التكامل الأخير نجد أنَّ:

$$\varphi_Y(t) = \mathbf{e}^{-\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \mathbf{e}^{-\frac{(y-\mathbf{i}t)^2}{2}}\, \mathbf{e}^{-\frac{t_2}{2}}\, dy = \mathbf{e}^{-\frac{t^2}{2}} \quad ; \forall t \in \mathbb{R}$$

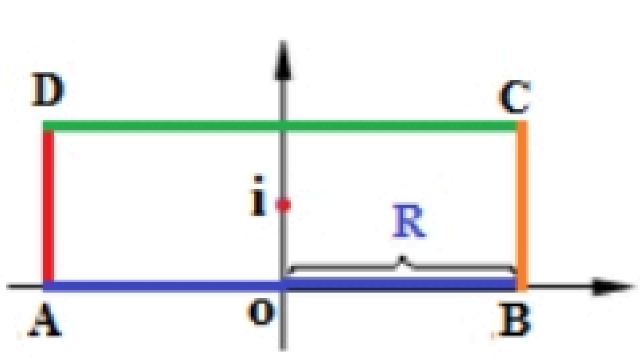
وبأخذ $X = \sigma\cdot Y + \mu$، و استخدام خواص الدالَّة المميِّزة، نجد الآتي:

$$\varphi_X(t) = \varphi_{\sigma\cdot Y + \mu}(t) = \mathbf{e}^{\mathbf{i}t\mu}\, \varphi_Y(\sigma t) = \mathbf{e}^{\mathbf{i}\mu t} \cdot \mathbf{e}^{-\frac{1}{2}\sigma^2 t^2} = \mathbf{e}^{\mathbf{i}\mu t - \frac{1}{2}\sigma^2 t^2}$$

في الواقع كان من الممكن الوصول إلى النتيجة السابقة نفسها من خلال حساب التكامل الآتي:

$$\varphi_X(t) = \mathbf{E}\, \mathbf{e}^{\mathbf{i}tX} = \int_{-\infty}^{\infty} \mathbf{e}^{\mathbf{i}t\,x} \cdot \frac{1}{\sigma\sqrt{2\pi}}\, \mathbf{e}^{-\frac{(x-\mu)^2}{2}}\, dx \quad ; \forall t \in \mathbb{R}$$

وذلك على المســـار المغلق ABCD الذي يحيط بالنقطة $\mathbf{i} = (0,1)$ (انظر الشكل الجانبي).




## (٣, ٤, ٧) نظريات شهيرة في دراسة الدوال المميِّزة

النظريات الآتية (التي سنقبلها دون برهان) تُعدّ من النظريات ا لمهمّة في إطار دراسة الدوال المميِّزة للمتغيِّرات العشوائية. إنَّ أول هذه النظريات تعطينا الشرط اللازم والكافي حتى تكون دالَّة عقدية معرفة على $\mathbb{R}$ هي دالَّة مميِّزة لمتغيِّر عشوائي، وتعرف هذه النظرية باسم نظرية بوخنر نسبة إلى الرياضياتي الأمريكي (هنغاري الأصل) بوخنر (1899-1982) Salomon Bochner.

### (١, ٣, ٤, ٧) نظرية بوخنر Bochner Theorem

إنَّ الشرط اللازم والكافي حتى تكون دالَّة عقدية $\varphi$ معرَّفة على $\mathbb{R}$ هي دالَّة مميِّزة لمتغيِّر عشوائي $X$ هو أنَّ تتحقَّق البنود الآتية:

١- من أجل كلّ $t \ni \mathbb{R}$ لدينا $|\varphi(t)| \geq 1$.

٢- الدالَّة $\varphi(t)$ مستمرَّة بانتظام على $\mathbb{R}$ .

٣- من أجل أيَّة قيم حقيقية كلّ $t_1$، $t_2$،...و $t_n$ من $\mathbb{R}$، و أيَّة قيم عقدية $z_1$، $z_2$،...و $z_n$ من $\mathbb{C}$ لدينا:

$$\sum_{j,k=1}^{n} \varphi(t_j - t_k)\, z_j\, \bar{z}_k \geq 0 \qquad [7,60]$$

وهذا البند الأخير يعني أنَّ الدالَّة $\varphi(t)$ غير سالبة التعيين.

من المسائل المهمّة أيضاً في نظرية الاحتمالات وتطبيقاتها (في نظرية المعاينة) موضوع حساب قيم التوزيع لمتغيِّر عشوائي باستخدام الدالَّة المميِّزة له، وهذا ما تقدمه لنا النظرية الآتية (سنقبلها دون برهان) والتي تعود إلى الرياضياتي الفرنسي ليفي Paul Pierre Lévy (1886-1971) التي قدَّمها عام 1925.

### (٢, ٣, ٤, ٧) نظرية ليفي Lévy Theorem

لتكن $X$ متغيِّر عشوائي فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، ولنفترض أنَّ $\varphi_X$ و $F_X$ هما الدالَّة المميِّزة ودالَّة التوزيع لمتغيِّر عشوائي $X$ على الترتيب. فإذا كانت الدالَّة $F_X$ مستمرَّة في جوار نقطة $x_o \in \mathbb{R}$ فإنَّه من أجل أيّ $\varepsilon > 0$ ستكون العلاقة الآتية مُحقَّقة:

$$F_X(x_o+\varepsilon) - F_X(x_o-\varepsilon) = \lim_{A\to\infty} \frac{1}{\pi} \int_{-A}^{A} \frac{\sin \varepsilon t}{t}\, e^{i t x_o}\, \varphi_X(t)\, dt \qquad [7,61]$$

وأكثر من ذلك، فإذا كان $\int_{-\infty}^{\infty} |\varphi_X(t)|\, dt < +\infty$ فإنَّ دالَّة الكثافة الاحتمالية $f_X$ ستكون موجودة في النقطة $x_o$ وتُعطى بالعلاقة الآتية:

$$f_X(x_o) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i t x_o}\, \varphi_X(t)\, dt \qquad [7,62]$$

لاحظ أنَّ هذه المبرهنة تظهر لنا وجود تقابل واحد لواحد بين دوال التوزيع الاحتمالية للمتغيِّرات العشوائية وبين الدوال المميِّزة لها، ولذلك تُعرف هذه النتيجة في الاحتمالات باسم نظرية الوحدانية.

أخيراً سنقدِّم إحدى المبرهنات الأخرى المتعلِّقة بالدوال المميِّزة للمتغيرات العشوائية والتي تُعْرَف باسم نظرية الاستمرَّار لـ ليفي ( وسنقبلها دون برهان أيضاً ).

### (٣, ٣, ٤, ٧) نظرية الاستمرار لـ ليفي Lévy's Continuity Theorem

لتكن $X_1$، $X_2$،...و $X_n$ و... متتالية متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ بدوال توزيع احتمالية $F_{X_1}$، $F_{X_2}$، ...، $F_{X_n}$ و... ودوال مميِّزة $\varphi_{X_1}$، $\varphi_{X_2}$، ...، $\varphi_{X_n}$ و... على الترتيب، فعندئذ:

١- إذا كانت المتتالية $\left(F_{X_n}\right)_{n\in\mathbb{N}}$ متقاربة من دالَّة توزيع احتمالية $F_X$ لمتغيِّر عشوائي ما $X$ فوق $[\Omega,\mathscr{A},P]$، فإنَّه سيكون لدينا ما يلي محقَّقاً:

$$\lim_{n\to\infty} \varphi_{X_n}(t) = \varphi_X(t) \qquad ; \forall t \in \mathbb{R}$$

٢- إذا كانت المتتالية $\left(\varphi_{X_n}\right)_{n\in\mathbb{N}}$ متقاربة من دالَّة مميِّزة $\varphi_X$ لمتغيِّر عشوائي ما $X$ فوق $[\Omega,\mathscr{A},P]$، وكانت هذه الدالَّة مستمرة في النقطة $t=0$ ، فإنَّه سيكون لدينا ما يلي محقَّقاً:

$$\lim_{n\to\infty} F_{X_n}(x) = F_X(x) \qquad ; \forall x \in \mathbb{R}$$

وذلك من أجل كلّ نقطة استمرار $x$ للدالَّة $F_X$ .

## (٧,٥) التوقع الشرطي

**Conditional Expectation**

إنَّ التوقع الشرطي يُعدّ من المفاهيم الضرورية في مجال العشوائيات لأنَّ مسائل عديدة تؤول في دراستها إلى موضوع التوقع الشرطي، لا بل أكثر من ذلك توجد قطاعات بحثية عديدة في نظرية الاحتمالات تقوم على مفهوم التوقع الشرطي بشكل أساسي ومنها على سبيل المثال لا الحصر **نظرية الحَكَمة** Martingale Theorem. إنَّ كلمة Martingale فرنسية تعني **"الحَكَمة"** وهي حديدة اللجام التي تكون في فم الفرس لإحكام جماحه، وسمّيت هذه النظرية بهذا الاسم لأنَّها بنيت في البدايات على رهانات سباق الخيل، ومدى تأثير الحَكَمة على نتائج السباق.

سنقوم فيما يلي بتوضيح مفهوم التوقع الشرطي، ولكن على سبيل التوضيح والتبسيط سنقوم بشرح لهذا المفهوم من أجل الحالة الخاصّة التي يكون فيها فضاء الحوادث الابتدائية $\Omega$ منتهياً، وكذلك المتغيِّرات العشوائية المستخدمة في هذه الدراسة هي متغيِّرات عشوائية متقطِّعة، ومن ثمَّ نعرض أهمَّ خصائص التوقع الشرطي من أجل هذه الحالة الخاصّة.

ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً مُعطى بحيث تكون $\Omega$ منتهية، ولنأخذ $\mathscr{Z}=\{Z_i \subseteq \Omega \ ; \ i \in \mathbb{N}_k\}$ تجزئة لفضاء الحوادث الابتدائية $\Omega$. عندئذ سنقوم بتعريف التوقع الشرطي وفقاً للمنهج الآتي:

**أولاً**: تعريف التوقع الشرطي لحادث $A$ بالنسبة إلى التجزئة $\mathscr{Z}$.

**ثانياً**: تعريف التوقع الشرطي لمتغيِّر عشوائي $X$ بالنسبة إلى التجزئة $\mathscr{Z}$.

### (٧,٥,١) التوقع الشرطي لحادث $A$ بالنسبة إلى تجزئة $\mathscr{Z}$

فيما يلي سنقوم بتوضيح مفهوم التوقع الشرطي لحادث $A$ بالنسبة إلى تجزئة $\mathscr{Z}$ ونبدأ ذلك بالتعريف الآتي.

#### (٧,٥,١,١) تعريف (التوقع الشرطي لحادث $A$ بالنسبة إلى تجزئة $\mathscr{Z}$)

ليكن $P(A \mid Z_i)$ هو الاحتمال الشرطي لحادث $A \in \mathscr{A}$ بالنسبة إلى ذرة التجزئة $Z_i$، فعندئذ يمكننا من خلال الاحتمالات الشرطية $\{P(A \mid Z_i)\ ;\ i \in \mathbb{N}_k\}$ أن نولِّد متغيِّراً عشوائياً $X_o$ فوق $[\Omega,\mathscr{A},P]$ وفقاً للعلاقة الآتية:

$$X_o(\omega) := \sum_{i=1}^{k} P(A \mid Z_i)\cdot \mathbf{I}_{Z_i}(\omega) \qquad [7,63]$$

وبهذا الشكل نلاحظ أنَّ كلّ تجزئة لـ $\Omega$ يوافقها متغيِّر عشوائي يُعرَّف وفقاً للعلاقة [7,63]، وأكثر من ذلك نلاحظ من هذه العلاقة الأخيرة أنَّ قيم $X_o$ على ذرات التجزئة $Z_i$ هي $P(A \mid Z_i)$ (مقادير ثابتة) من أجل كلّ $i \in \mathbb{N}_k$، ولكي نميِّز العلاقة الخاصّة بالمتغيِّر العشوائي $X$ مع التجزئة $\mathscr{Z}$ سوف نكتب:

$$X_o(\omega) \triangleq P(A \mid \mathscr{Z})(\omega) \qquad ;\forall \omega \in \Omega \qquad [7,64]$$

علماً أنَّ العلاقة السابقة تُدعى **الاحتمال الشرطي للحادث** $A \in \mathscr{A}$ **بالنسبة إلى التجزئة** $\mathscr{Z}$.

#### (٧,٥,١,٢) خصائص الاحتمال الشرطي بالنسبة إلى تجزئة

١- من أجل أي حادثين متنافيين $A, B \in \mathscr{A}$ تكون العلاقة الآتية مُحقَّقةً:

$$P(A+B \mid \mathscr{Z}) = P(A \mid \mathscr{Z}) + P(B \mid \mathscr{Z}) \qquad [7,65]$$

٢- في الحالة الخاصة إذا كانت $\mathscr{Z}$ هي التجزئة المبتذلة (أي إنَّ $\mathscr{Z}=\{\Omega\}$)، فعندئذ ستكون العلاقة الآتية مُحقَّقةً:

$$P\left(A \mid \mathcal{Z}\right)+P\left(A\right) \qquad [7,66]$$

وذلك لأنَّ $\mathbf{I}_{\Omega}(\omega)=1$ من أجل كلّ $\omega\in\Omega$، ومن ثمَّ يكون لدينا:

$$P\left(A \mid \mathcal{Z}\right)(\omega)=\sum_{i=1}^{k} P(A \mid Z_i)\cdot\mathbf{I}_{Z_i}(\omega)=P(A \mid \Omega)\cdot\mathbf{I}_{\Omega}(\omega)=P(A \mid \Omega)=\frac{P(A\cap\Omega)}{P(\Omega)}=P(A)$$

لاحظ أنَّ تعريف الاحتمال الشرطي لحادث بالنسبة إلى تجزئة كمتغيِّر عشوائي يعطينا الإمكانية في التحدث عن قيمته المتوقعة، ومن ثمَّ يمكننا استغلال هذا الطرح من أجل تقديم صيغة الاحتمال التام، ويتمّ ذلك على النحو الآتي:

ليكن $A\in\mathcal{A}$ كيفياً، فعندئذ يمكننا أن نكتب الآتي (**صيغة الاحتمال التام**):

$$P(A)=\mathbf{E}\left[P\left(A \mid \mathcal{Z}\right)\right] \qquad [7,67]$$

وذلك لأنَّه من تعريف $X_o$ يمكننا أن نكتب الآتي:

$$\mathbf{E}\left[X_o(\omega)\right]=\mathbf{E}\left[P\left(A \mid \mathcal{Z}\right)\right](\omega)=\sum_{i=1}^{k}P(A \mid Z_i)\cdot\mathbf{E}\left[\mathbf{I}_{Z_i}(\omega)\right]$$

ولكن نعلم أنَّ $\mathbf{E}\left[\mathbf{I}_{Z_i}(\omega)\right]=P(Z_i)$، ومن ثمَّ يصبح لدينا:

$$\mathbf{E}\left[P\left(A \mid \mathcal{Z}\right)\right](\omega)=\sum_{i=1}^{k}P(A\cap Z_i)=P(A)$$

وبهذا نكون قد حصلنا على صيغة الاحتمال التام.

لنقم الآن بتعريف التجزئة المولَّدة من متغيِّر عشوائي على النحو الآتي.

### (٧,٥,١,٣) تعريف (التجزئة المولَّدة من متغيِّر عشوائي The Generated Partition by a Random Variable)

ليكن $Y$ متغيِّراً عشوائياً متقطِّعاً فوق $[\Omega,\mathcal{A},P]$ بمجموعة قيم $\mathbf{Y}=\{y_j\ ;\ j\in\mathbb{N}_k\}$ واحتمالات $P(Y=y_j)>0$ من أجل كلّ $j\in\mathbb{N}_k$، ومُعرَّفاً بالعلاقة الآتية:

$$Y(\omega):=\sum_{j=1}^{k} y_j\cdot\mathbf{I}_{Z_j}(\omega) \qquad [7,68]$$

وبملاحظة أنَّ $Y$ يأخذ قيماً ثابتةً على الذرات $Z_j$، فإنَّه سيكون لهذه الذرات العرض الآتي:

$$Z_j=\left\{\omega\in\Omega\ ;Y(\omega)=y_j\right\} \qquad ;\ j\in\mathbb{N}_k \qquad [7,69]$$

إنَّ التجزئة $\mathcal{Z}$ التي ذراتها مُعطاة وفقاً للعلاقة [7,69] تُسمى **التجزئة المولَّدة بـ $Y$** ويُرمز لها بـ $\mathcal{Z}_Y$ أي إنَّه لدينا:

$$\mathcal{Z}_Y=\left\{Z_j : Z_j=\{\omega\in\Omega\ ;Y(\omega)=y_j\}\ ;\forall j\in\mathbb{N}_k\right\} \qquad [7,70]$$

### (٧,٥,١,٤) ملاحظات

١- يُشار عادة إلى الاحتمال الشرطي $P\left(A \mid \mathcal{Z}_Y\right)$ بالرمز $P(A \mid Y)$ أو بالشكل $P(A \mid Y)(\omega)$ لتبيان أنَّه متغيِّر عشوائي، ويُقال له الاحتمال الشرطي للحادث $A$ بالنسبة إلى المتغيِّر العشوائي $Y$.

٢- سنتفق على أنَّ الكتابة $P\left(A \mid Y=y_j\right)$ تعني الاحتمال الشرطي $P(A \mid Z_j)$ علماً أنَّ $Z_j$ مُعطاة وفقاً للعلاقة [7,69] من أجل كلّ $j\in\mathbb{N}_k$.

٣- إنَّ المتغيِّر العشوائي $P\left(A \mid \mathfrak{Z}_Y\right)$ قيوساً بالنسبة إلى الـ σ- جبر $\mathfrak{S} = \sigma\left(\mathfrak{Z}\right)$ (الـ σ- جبر المولَّد من التجزئة $\mathfrak{Z}$)، ولذلك يُرمز له بـ $P\left(A \mid \mathfrak{S}\right)$ أيضاً.

بهذا الشكل يكون مفهوم الاحتمال الشرطي لحادث بالنسبة إلى σ- جبر مبرراً وله معنى.

**(٧, ٥, ١, ٥) تعريف (التجزئة المولَّدة من عدد منته من المتغيِّرات العشوائية)**

ليكن $Y_1$، $Y_2$،...و $Y_m$ متغيِّرات عشّوائيّة متقطّعةً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، وسنفترض أنَّ مجموعات قيمها على الترتيب هي $\mathbf{Y}_\ell = \left(y_{\ell k_s} \;\; ; s \in \mathbb{N}_{n_\ell}\right)$ من أجل كلّ $\ell$ في $\mathbb{N}_m$، وكذلك سنفترض أنَّ:

$$P\left(Y_\ell = y_{\ell k_s}\right) > 0 \quad ; s \in \mathbb{N}_{n_\ell} \;,\; \ell \in \mathbb{N}_m \qquad [7,71]$$

ولنعرِّف المجموعات الآتية:

$$Z_{y_{1k_1}, y_{2k_2}, \dots, y_{mk_m}} := \left(\omega \in \Omega : Y_\ell(\omega) = y_{\ell k_\ell} \;\; ; k_\ell \in \mathbb{N}_{n_\ell} \,,\, \ell \in \mathbb{N}_m\right) \qquad [7,72]$$

عندئذ تُدعى أسرة الحوادث (المجموعات):

$$\mathfrak{Z}_{Y_1, Y_2, \dots, Y_m} := \left(Z_{y_{1k_1}, y_{2k_2}, \dots, y_{mk_m}} \;\; ; \forall k_\ell \in \mathbb{N}_{n_\ell} \,,\, \ell \in \mathbb{N}_m\right)$$

التجزئة المولَّدة من المتغيِّرات العشوائية $Y_1$، $Y_2$،...و $Y_m$.

**(٧, ٥, ١, ٦) ملاحظات**

١- يُشار عادة إلى الاحتمال الشرطي $P\left(A \mid \mathfrak{Z}_{Y_1}, \mathfrak{Z}_{Y_2}, \dots, \mathfrak{Z}_{Y_m}\right)$ بالرمز $P\left(A \mid Y_1, Y_2, \dots, Y_m\right)(\omega)$ لتبيان أنَّه متغيِّر عشوائي، ويُدعى الاحتمال الشرطي للحادث $A$ بالنسبة إلى المتغيِّرات العشوائية $Y_1$، $Y_2$،...و $Y_m$.

٢- سنتفق على أنَّ الاصطلاح الآتي:

$$P\left(A \mid Z_{y_{1k_1}, y_{2k_2}, \dots, y_{mk_m}}\right) \triangleq P\left(A \mid Y_1 = y_{1k_1} \,,\, Y_2 = y_{2k_2} \,, \dots, Y_m = y_{mk_m}\right)$$

علماً أنَّ المجموعة $Z_{y_{1k_1}, y_{2k_2}, \dots, y_{mk_m}}$ مُعطاة وفقاً للعلاقة [7,72] .

## (٧, ٥, ٢) التوقع الشرطي لمتغيِّر عشوائي بالنسبة إلى تجزئة

## The Conditional Expectation of a R.V. with respect to a Partition

سنقوم بتوضيح مفهوم التوقع الشرطي لمتغيِّر عشوائي $X$ بالنسبة إلى تجزئة $\mathfrak{Z}$ مُعطاة.

**(٧, ٥, ٢, ١) تعريف (التوقع الشرطي لمتغيِّر عشوائي بالنسبة إلى تجزئة)**

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ بمجموعة قيم $\mathbf{X} = \{x_1, x_2, \dots, x_n\}$ (أي إنَّ $X$ متغيِّر عشوائي بسيط)، فعندئذ يمكن عرضه على النحو الآتي:

$$X = \sum_{i=1}^{n} x_i \, \mathbf{I}_{A_i} \qquad ; \forall i \in \mathbb{N}_n$$

علماً أنَّ الحوادث $A_i$ لها العرض الآتي:

$$A_i = \{\omega \in \Omega \;;\; X(\omega) = x_i\} \qquad ;\forall i \in \mathbb{N}_n$$

ومن ثمَّ نجد أنَّ التوقع الرياضياتي لـ $X$ يساوي:

$$\mathbf{E}\, X = \sum_{i=1}^{n} x_i \cdot P(A_i) \qquad [7,73]$$

وذلك لأنَّ:

$$\mathbf{E}\, X = \sum_{i=1}^{n} x_i \cdot P(X = x_i) = \sum_{i=1}^{n} x_i \cdot \mathbf{E}\left(\mathbf{I}_{A_i}\right) = \sum_{i=1}^{n} x_i \cdot P(A_i)$$

والآن لو أخذنا $\mathcal{Z} = \{Z_j \subseteq \Omega : j \in \mathbb{N}_k\}$ تجزئة لـ $\Omega$، فإنَّه يمكننا باستخدام الاحتمالات الشرطية $P(A_i \mid \mathcal{Z})$ لكل $i \ni \mathbb{N}_n$ أنَّ نعرِّف التوقع الشرطي للمتغيِّر العشوائي $X$ بالنسبة إلى التجزئة $\mathcal{Z}$، حيث يُرمز له بـ $\mathbf{E}(X \mid \mathcal{Z})(\omega)$ لتبيان أنَّ هذا التوقع الشرطي هو متغيِّر عشوائي، حيث يوضع:

$$\mathbf{E}(X \mid \mathcal{Z})(\omega) = \mathbf{E} \sum_{i=1}^{n} x_i \cdot P\left(A_i \mid \mathcal{Z}\right)(\omega) \qquad [7,74]$$

### (٢, ٢, ٥, ٧) ملاحظات

١- وفقاً لهذا التعريف يكون التوقع الشرطي للمتغيِّر العشوائي $X$ بالنسبة إلى تجزئة $\mathcal{Z}$ هو متغيِّر عشوائي فوق $[\Omega, \mathcal{A}, P]$، وقيمته من أجل أيّ حادث ابتدائي $\omega \ni Z_j$ يساوي $\sum_{i=1}^{n} x_i\, P(A_i \mid Z_j)$. إنَّ هذه الملاحظة توضِّح لنا أنَّه كان بالإمكان التوصل إلى تعريف التوقع الشرطي لمتغيِّر عشوائي $X$ بالنسبة إلى تجزئة $\mathcal{Z}$ من خلال تعريف التوقع الشرطي $\mathbf{E}(X \mid Z_j)$ وذلك باستخدام العلاقة الآتية:

$$\mathbf{E}\left(X \cdot \mathbf{I}_{Z_j}\right) = \sum_{i=1}^{n} x_i\, P(A_i\, Z_j)$$

والتي تعطينا العلاقة الآتية:

$$\mathbf{E}\left(X \mid Z_j\right) = \sum_{i=1}^{n} x_i\, P(A_i\, Z_j) \qquad [7,75]$$

وبعد ذلك لو وضعنا بالتعريف ما يلي:

$$\mathbf{E}(X \mid \mathcal{Z}) := \sum_{j=1}^{k} \mathbf{E}\left(X \mid Z_j\right) \cdot \mathbf{I}_{Z_j} \qquad [7,76]$$

سيصبح لدينا ما يلي مُحقَّقاً:

$$\mathbf{E}(X \mid \mathcal{Z})(\omega) = \sum_{j=1}^{k} \mathbf{E}\left(X \mid Z_j\right) \cdot \mathbf{I}_{Z_j}(\omega) \qquad ;\forall \omega \in \Omega \qquad [7,77]$$

٢- يتبيَّن لنا من العروض السابقة أنَّ $\mathbf{E}(X \mid \mathcal{Z})$ و $\mathbf{E}(X \mid Z_j)$ غير متعلِّقة بنوعية العرض للمتغيِّر العشوائي $X$.

٣- يمكننا كتابة $\mathbf{E}(X \mid Z_j)$ على النحو الآتي أيضاً:

$$\mathbf{E}\left(X \mid Z_j\right) = \frac{\mathbf{E}\left(X \cdot \mathbf{I}_{Z_j}\right)}{P(Z_j)} \qquad ;\forall j \in \mathbb{N}_k$$

### (٧,٥,٢,٣) تعريف (التجزئة التي أنْعم من تجزئة أخرى Partition Smoother Than Another Partition)

لتكن $\mathcal{Z}_1$ و $\mathcal{Z}_2$ تجزئتين مختلفتين لـ $\Omega$، فعندئذ يُقال إنَّ التجزئة $\mathcal{Z}_1$ أنْعم من التجزئة $\mathcal{Z}_2$ إذا كان من أجل كلّ ذرة $A \in \mathcal{Z}_1$ توجد ذرة $B \in \mathcal{Z}_2$ بحيث تكون $A \subseteq B$، وسنرمز لذلك بـ $\mathcal{Z}_1 \gg \mathcal{Z}_2$.

### (٧,٥,٢,٤) تعريف (المتغيِّر العشوائي القيوس بالنسبة إلى تجزئة)
### Random Variable, that is measurable with respect to a Partition

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$، و $\mathcal{Z}_X$ تجزئة مولَّدة من $X$، وبفرض أنَّ $\mathcal{Z}$ تجزئة ما لـ $\Omega$. عندئذ يُقال إنَّ المتغيِّر العشوائي $X$ **قيوس بالنسبة إلى التجزئة** $\mathcal{Z}$ (**أو يُقال $X$ هو $\mathcal{Z}$-قيوس** ($X$ is $\mathcal{Z}$ - measurable)) إذا كانت $\mathcal{Z}_X \gg \mathcal{Z}$.

لاحظ هنا أنَّه إذا كان $X$ قيوساً بالنسبة إلى التجزئة $\mathcal{Z} = \{Z_j \subseteq \Omega : j \in \mathbb{N}_k\}$ فإنَّ المتغيِّر العشوائي $X$ سيكون قابلاً للعرض من خلال العلاقة الآتية:

$$X(\omega)= \sum_{i=1}^{k} x_i \cdot \mathbf{I}_{Z_i}(\omega) \qquad [7,78]$$

ومن هذا يتضِّح لنا أنَّ المتغيِّر العشوائي $X$ سيكون قيوساً بالنسبة إلى التجزئة $\mathcal{Z}$ إذا كان $X$ يأخذ قيماً ثابتةً على ذرات التجزئة $\mathcal{Z}$ (**أي إنَّ لـ $X$ شكل دالَّة درجية على ذرات التجزئة** $\mathcal{Z}$)، وكمثال على ذلك لو أخذنا $\mathcal{Z} = \{\Omega\}$ هي التجزئة المبتذلة، فعندئذ سيكون $X$ قيوساً بالنسبة إلى التجزئة $\mathcal{Z}$ **إذا وفقط إذا** كان $X$ ثابتاً على $\Omega$ باحتمال يساوي الواحد، أي إنَّ $X$ خاضع للتوزيع الوحيد النقطة.

### (٧,٥,٢,٥) نتيجة

إنَّ أي متغيِّر عشوائي $X$ فوق فضاء احتمالي $[\Omega,\mathcal{A},P]$ سيكون قيوساً بالنسبة إلى التجزئة المولَّدة منه $\mathcal{Z}_X$ وذلك لأنَّه يأخذ قيماً ثابتةً على ذراتها.

### (٧,٥,٢,٦) ملاحظات

١- عندما تكون التجزئة $\mathcal{Z}$ مولَّدة من متغيِّرات عشوائية $Y_1$، $Y_2$،...و $Y_m$، أي إنَّ $\mathcal{Z}=\mathcal{Z}_{Y_1,Y_2,\dots,Y_m}$، فعندئذ يُرمز للتوقع الشرطي لـ $X$ بالنسبة إلى التجزئة $\mathcal{Z}_{Y_1,Y_2,\dots,Y_m}$ بالرمز $\mathbf{E}(X \mid Y_1, Y_2,\dots,Y_m)(\omega)$ ويُطلق عليه اسم **التوقع الشرطي لـ $X$ بالنسبة إلى المتغيِّرات العشوائية** $Y_1$، $Y_2$،...و $Y_m$.

٢- من تعريف $\mathbf{E}(X \mid Y)$ ينتج لدينا أنَّه إذا كان $X$ و $Y$ مستقلَّان بعضهما عن بعض فإنَّ العلاقة الآتية ستكون مُحقَّقة:

$$\mathbf{E}(X \mid Y) = \mathbf{E}X \qquad [7,79]$$

وذلك لأنَّه لدينا:

$$\mathbf{E}(X \mid Y) = \mathbf{E}(X \mid \mathcal{Z}_Y)(\omega) = \sum_k \mathbf{E}(X \mid Z_k)\cdot \mathbf{I}_{Z_k}(\omega)$$

وبما أنَّ $\sum_{k=1}^{m} \mathbf{I}_{Z_k}(\omega) = 1$، وكذلك لدينا:

$$\mathbf{E}(X \mid Z_k) = \sum_i x_i \cdot P(A_i \mid Z_k) = \mathbf{E}X$$

وبهذا تنتج لدينا صحة العلاقة [7,79] .

٣- لدينا $\mathbf{E}(X \mid X) = X$، وذلك لأنَّ:

$$\mathbf{E}(X \mid X) = \mathbf{E}(X \mid \mathcal{Z}_X)(\omega) = \sum_k \mathbf{E}(X \mid Z_k)\cdot \mathbf{I}_{Z_k}(\omega) = \sum_k \sum_\ell x_\ell P(Z_\ell \mid Z_k)\mathbf{I}_{Z_k}(\omega)$$

ولكن لدينا:

$$P\left(A_{\ell}|A_{k}\right)=\begin{cases}1 & for \quad \ell=k\\ \\ 0 & for \quad \ell\neq k\end{cases}$$

ومنه ينتج لدينا:

$$\sum_{k}\ \sum_{\ell}x_{\ell}P\left(Z_{\ell}\mid Z_{k}\right)\mathbf{I}_{Z_{k}}(\omega)=\sum_{k}x_{k}\cdot\mathbf{I}_{Z_{k}}(\omega)=X(\omega)$$

وبهذا يتمّ الإثبات.

سنقدِّم فيما يلي بعض خصائص التوقع الشرطي وفقاً للطرح المقدَّم سابقاً، وهذه الخصائص سنراها ذاتها لدى تقديم العرض المعمَّم للتوقع الشرطي، وأمّا الإثباتات لهذه الخصائص فسوف تترك كتمارين للقارئ.

**(٧,٢,٥,٧) مبرهنة (خصائص التوقع الشرطي بالنسبة إلى تجزئة)**

ليكن $X$ و$Y$ متغيِّرين عشوائيين فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، ولنأخذ $\mathscr{Z}=\left\{Z_i\subseteq\Omega:\ i\in\mathbb{N}_k\right\}$ تجزئة لـ $\Omega$، فعندئذ:

١- بفرض أنَّ $a$ و $b$ أعداد حقيقية مع $a\neq 0$، فإنَّه سيكون لدينا:

$$\mathbf{E}\left(aX+bY\mid\mathscr{Z}\right)=a\cdot\mathbf{E}\left(X\mid\mathscr{Z}\right)+b\cdot\mathbf{E}\left(Y\mid\mathscr{Z}\right)$$

٢- من أجل التجزئة المبتذلة $\mathscr{Z}=\{\Omega\}$ لدينا:

$$\mathbf{E}\left(X\mid\{\Omega\}\right)=\mathbf{E}X$$

٣- إذا كان $X=c$ على $\Omega$ باحتمال يساوي الواحد، فإنَّه سيكون لدينا:

$$\mathbf{E}\left(X\mid\mathscr{Z}\right)=c$$

٤- من أجل أيّ حادث $A\in\mathscr{A}$ يكون لدينا:

$$\mathbf{E}\left(\mathbf{I}_A\mid\mathscr{Z}\right)(\omega)=P\left(A\mid\mathscr{Z}\right)(\omega) \qquad [7,80]$$

ومن هذه العلاقة يتبيَّن لنا أنَّ خصائص الاحتمال الشرطي يمكن الحصول عليها مباشرة من خصائص التوقع الشرطي.

٥- لدينا العلاقة الآتية محقَّقةً دوماً:

$$\mathbf{E}\left[\mathbf{E}\left(X\mid\mathscr{Z}\right)\right]=\mathbf{E}X \qquad [7,81]$$

إنَّ هذه الخاصيَّة تُعمِّم لنا صيغة الاحتمال التام المعطى من خلال العلاقة [7,67]، وذلك لأنَّه إذا أخذنا الحالة الخاصة $X=\mathbf{I}_A$ فإنَّ [7,67] ستنتج من العلاقة السابقة [7,81].

٦- يكون $X$ قيوساً بالنسبة إلى التجزئة $\mathscr{Z}$ إذا كان $X$ يأخذ قيماً ثابتة على ذرات هذه التجزئة، وهذه الحقيقة تنتج مباشرة من العلاقة [7,78].

٧- إذا كان $X$ قيوساً بالنسبة إلى التجزئة $\mathscr{Z}$، فعندئذ ستكون العلاقة الآتية مُحقَّقة:

$$\mathbf{E}\left(X\cdot Y\mid\mathscr{Z}\right)=X\cdot\mathbf{E}\left(Y\mid\mathscr{Z}\right) \qquad [7,82]$$

وبشكل خاص عندما $Y = 1$ باحتمال يساوي الواحد، فإنَّه سينتج لدينا:

$$\mathbf{E}\left(X \mid \mathcal{Z}\right) = X \qquad \& \qquad \mathbf{E}\left(X \mid \mathcal{Z}_X\right) = X$$

٨- لتكن $\mathcal{Z}_1$ و $\mathcal{Z}_2$ تجزئتين لـ $\Omega$ بحيث إنَّ $\mathcal{Z}_1$ (( $\mathcal{Z}_2$ ، فعندئذ يكون لدينا:

$$\mathbf{E}\left[\ \mathbf{E}\left(X \mid \mathcal{Z}_2\right) \middle| \ \mathcal{Z}_1\right] = \mathbf{E}\left(X \mid \mathcal{Z}_1\right) \qquad [7,83]$$

### (٧,٥,٢,٨) نتائج:

١- من تعريف $\mathbf{E}(X \mid Y)$ ينتج أنَّه إذا كان $X$ و$Y$ مستقلَّين فإنَّ العلاقتين الآتيتين ستكونان مُحقَّقتين:

$$\textbf{a)}\ \ \mathbf{E}\left(X \mid Y\right) = X \qquad \& \qquad \textbf{b)}\ \ \mathbf{E}\left(Y \mid X\right) = Y \qquad [7,84]$$

٢- إنَّ الخاصة (**a**) من [7,84] تسمح لنا بالتعميم الآتي:

إذا كان $X$ مستقلًا عن التجزئة $\mathcal{Z}$، أي إنَّه من أجل كلّ $Z_j \in \mathcal{Z}$ يكون $X$ و $\mathbf{I}_{Z_j}$ مستقلَّين بعضهما عن بعض، فعندئذ سيكون لدينا:

$$\mathbf{E}\left(X \mid \mathcal{Z}\right) = \mathbf{E}X \qquad [7,85]$$

٣- من العلاقة السابقة [7,85] نحصل بشكل خاص على العلاقة الآتية:

$$\mathbf{E}\left[\mathbf{E}\left(X \mid Y_1, Y_2\right) \middle| \ Y_1\right] = \mathbf{E}\left(X \mid Y_1\right) \qquad [7,86]$$

### (٧,٥,٢,٩) ملاحظات

١- نعلم أنَّه في حال أنَّ $\Omega$ منتهية فإنَّ مفهومي الجبر والـ $\sigma$- **جبر** فوق $\Omega$ سيكونان متطابقان، ولهذا فإنَّ التوقع الشرطي لمتغيِّر عشوائي $X$ بالنسبة إلى جبر ما $\mathcal{B}$ فوق $\Omega$ سيتطابق مع مفهوم التوقع الشرطي لمتغيِّر عشوائي $X$ بالنسبة إلى تجزئة **مُولِّدة** للجبر $\mathcal{B}$، فلو كان $\mathcal{B} = \alpha\left(\mathcal{Z}\right)$ هو الجبر المولَّد من التجزئة $\mathcal{Z}$ فإنَّه يُفهم من $\mathbf{E}\left(X \mid \mathcal{B}\right)$ على أنَّه بالتعريف هو $\mathbf{E}\left(X \mid \mathcal{Z}\right)$.

٢- في مجال الإحصاء الرياضياتي تُعرَّف الدالَّة:

$$R\left(y_j\right) = \mathbf{E}\left(X \ \middle| \ Y = y_j\right) \qquad [7,87]$$

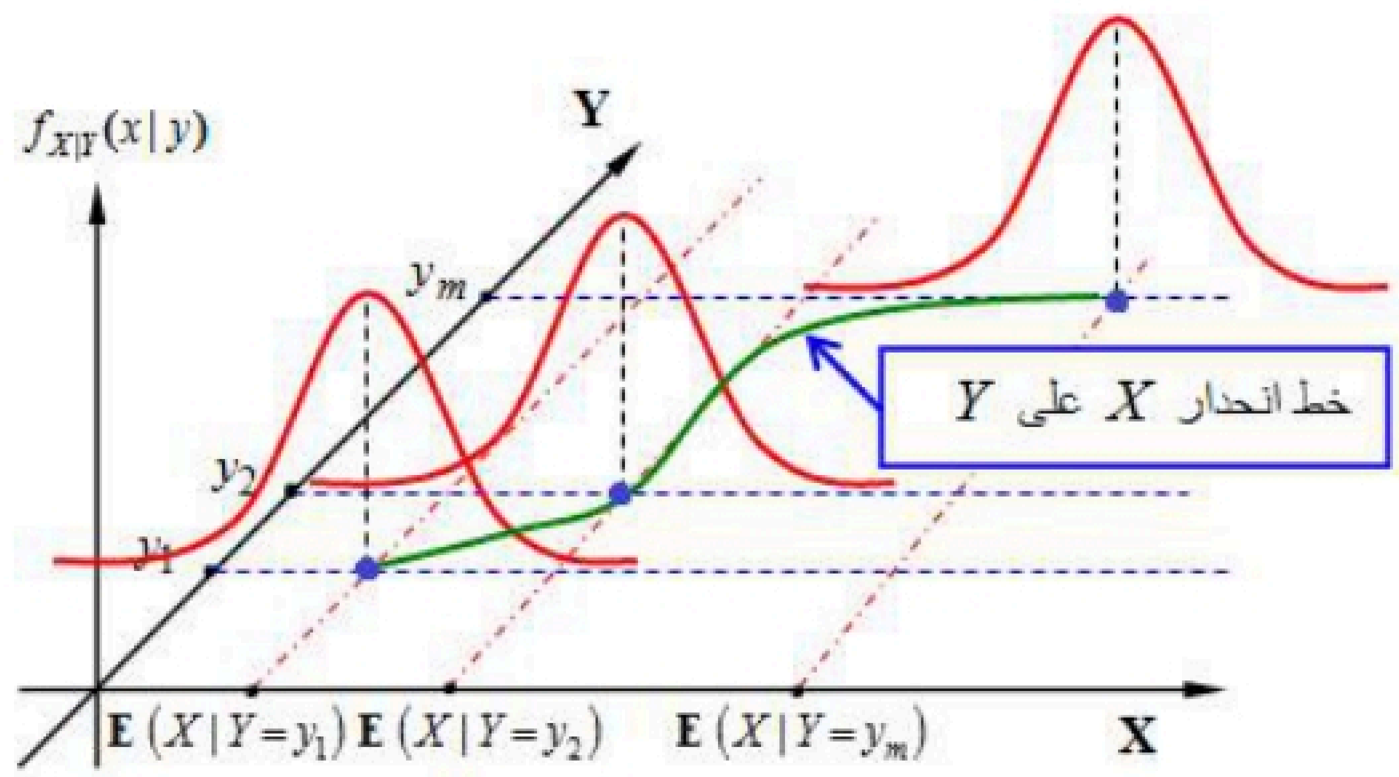
باسم **معادلة انحدار** Regression Equation **المتغيِّر** $X$ **على** $Y$ ، وأمَّا المنحنى النّاتج عن المجموعة $\left\{R\left(y_j\right) ; j \in J\right\}$ فإنَّه يُعرَّف باسم **خط انحدار** Regression Line **المتغيِّر** $X$ **على** $Y$ ، وبالمثل يُطلق على الدالَّة:

$$R\left(x_i\right) = \mathbf{E}\left(Y \ \middle| \ X = x_i\right) \qquad [7,88]$$

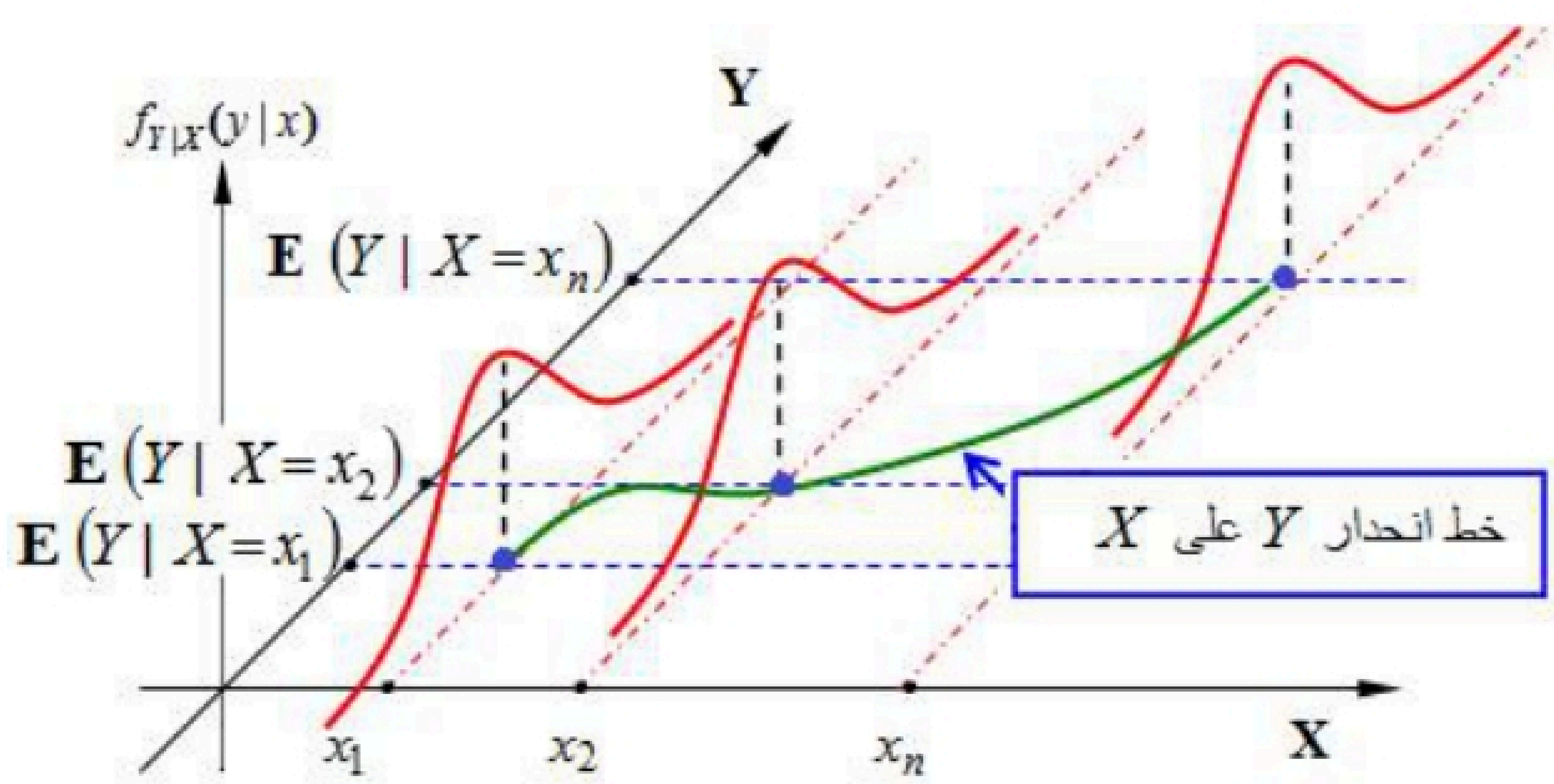
اسم **معادلة انحدار المتغيِّر** $Y$ **على** $X$ ، وأمّا المنحنى النّاتج عن المجموعة $\left\{R\left(x_i\right) ; i \in I\right\}$ فإنَّه يُدعى **خط انحدار** $Y$ **على** $X$.

أخيراً نشير هنا إلى أنَّ المعادلتين [7,87] و [7,88] تُدعيان **معادلتي الانحدار** للمسألة الإحصائية التي قيد الدراسة.

في الواقع يُنظر إلى كلّ من المتغيِّرين العشوائيين $X$ و$Y$ على أنَّهما مركَّبات متَّجه عشوائي $\mathbb{X}_2 = (X, Y)$ راصد للمحوظة خاصّة بالتجربة التي قيد الدراسة، فعلى سبيل المثال لو كان للمتَّجه العشوائي $\mathbb{X}_2 = (X, Y)$ توزيع طبيعي (ثنائي البُعد) حيث يبرهن على أنَّ التوزيعين الهامشيين لكلٍّ من $X$ و$Y$ سيكون لهما توزيع طبيعي أيضاً، فعندئذ سيكون لبيان الدالَّتين الممثِّلتين لمعادلتي الانحدار خطان واقعان في المستوي **XoY** والمشار إليهما في الشكلين الآتيين (١,٧. أ) و (١,٧. ب).




الشكل (١,٧.أ)




الشكل (١,٧.ب)

**(١٠, ٢, ٥, ٧) مثال:**

ليكن $X$ و$Y$ متغيِّرين عشوائيين مستقلَّين متطابقي التوزيع ويأخذان القيمتين 1 و0 باحتمال $0<p<1$ و $q=1-p$ على الترتيب، ولنقم بما يلي:

أ– حساب الاحتمال الشرطي $P(X+Y=k \mid Y)$ وذلك من أجل أي $k$ ∍ $\{0, 1, 2\}$.

ب– حساب التوقع الشرطي $\mathbf{E}(X+Y \mid Y)$.

**الحـل:** من أجل ذلك نعلم أنَّه إذا كان $X$ و$Y$ متغيِّرين عشوائيين مستقلَّين ويأخذان القيمتين $x$ و$y$ على الترتيب، فإنَّه سيكون لدينا ما يلي مُحَقَّقاً:

$$P(X+Y=z \mid Y=y) = P(X+y=z)$$

وذلك لأنَّه باستخدام تعريف الاحتمال الشرطي ومفهوم الاستقلال نجد الآتي:

$$P(X+Y=z \mid Y=y) = \frac{P(X+Y=z\ ,\ Y=y)}{P(Y=y)} = \frac{P(X+y=z\ ,Y=y)}{P(Y=y)}$$
$$= \frac{P(X+y=z)\cdot P(Y=y)}{P(Y=y)} = P(X+y=z)$$

ومن ثمَّ باستخدام هذه العلاقة الأخيرة نجد أنَّ:

$$P\left(X+Y=k \mid Y\right)=P\left(X+Y=k \mid Y=0\right)\cdot \mathbf{I}_{\{Y=0\}}(\omega)+P\left(X+Y=k \mid Y=1\right)\cdot \mathbf{I}_{\{Y=1\}}(\omega)$$

وهذه العلاقة الأخير يمكننا كتابتها بالشكل الآتي:

$$P\left(X+Y=k \mid Y\right)=P\left(X=k\right)\cdot \mathbf{I}_{\{Y=0\}}(\omega)+P\left(X=k-1\right)\cdot \mathbf{I}_{\{Y=1\}}(\omega)$$

والتي نجد منها:

$$P\left(X+Y=k \mid Y\right)=\begin{cases} q\cdot \mathbf{I}_{\{Y=0\}}(\omega) & for\ k=0 \\ p\cdot \mathbf{I}_{\{Y=0\}}(\omega)+q\cdot \mathbf{I}_{\{Y=1\}}(\omega) & for\ k=1 \\ p\cdot \mathbf{I}_{\{Y=1\}}(\omega) & for\ k=2 \end{cases}$$

وهكذا ينتج لدينا الآتي:

$$P\left(X+Y=k \mid Y\right)=\begin{cases} q\cdot(1-Y) & for\ k=0 \\ p\cdot(1-Y)+q\cdot Y & for\ k=1 \\ p\cdot Y & for\ k=2 \end{cases}$$

ب – من العلاقتين **(a)** و **(b)** من [7,84]، وبسبب استقلال $X$ عن $Y$ ينتج لدينا:

$$\mathbf{E}\left(X+Y \mid Y\right)=\mathbf{E}\left(X \mid Y\right)+\mathbf{E}\left(Y \mid Y\right)=\mathbf{E}X+Y=p+Y$$

وكان من الممكن الحصول على هذه النتيجة اعتماداً على تعريف التوقع الشرطي لمتغيِّر عشوائي أيضاً، وبهذا يتمّ الحل.

**هذا ما تيسَّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل**

# تمارين الفصل السابع

١- لنأخذ تجربة قذف قطعة نقود متوازنة لثلاث مرات متتالية، وليكن $X$، $Y$ و$Z$ ثلاثة متغيِّرات عشوائية فوق الفضاء الاحتمالي لهذه التجربة العشوائية، ومعرّفة كما يلي:

$$X(\omega) := \begin{cases} -1 & for \quad \omega = \{HHH\} \\ 1 & for \quad \omega = \{TTT\} \\ 0 & otherwise \end{cases}$$

$$Y(\omega) := \begin{cases} 1 & for \ \ \omega = \{TTT, HHH\} \\ -1 & for \ \ \omega = \{THH, HTH, HHT\} \\ 0 & for \ \ \omega = \{THT, TTH, HTT\} \end{cases}$$

$$Z(\omega) := \begin{cases} -1 & for \ \ \omega = \{HHH, HHT, HTH, THH\} \\ 1 & for \ \ \omega = \{TTT, TTH, THT, HTT\} \end{cases}$$

والمطلوب ما يلي:

أ- بين إنَّ كانت هذه المتغيِّرات العشوائية مستقلّة أم لا؟

ب- احسب التباين للمتغيِّر العشوائي $X$ وفقاً لتعريف التباين.

ج- عيِّن الدالَّة المولِّدة لعزوم المتغيِّر العشوائي $Y$، ومن ثمَّ استخدمها من أجل حساب التباين لهذا المتغيِّر العشوائي.

د- عيِّن الدالَّة المميِّزة للمتغيِّر العشوائي $Z$، ومن ثمَّ استخدمها من أجل حساب التباين لهذا المتغيِّر العشوائي.

هـ- احسب قيمة مُعامل الارتباط الخطي بين $X$ و$Z$، ومن ثمَّ ناقش قيمته.

٢- لنأخذ تجربة إلقاء حجري نرد متمايزين ومتوازنين لمرّة واحدة فقط، وليكن $X$ و$Y$ متغيِّرين عشوائيين فوق الفضاء الاحتمالي لهذه التجربة معرّفين من خلال العلاقتين الآتيتين:

$$X((i,j)) := \begin{cases} -2 & for \ \ i+j \le 4 \\ 0 & for \ \ 4 < i+j \le 8 \\ 6 & for \ \ i+j > 8 \end{cases}$$

$$Y((i,j)) := \begin{cases} 0 & for \ i+j \le 7 \\ 1 & for \ i+j > 7 \end{cases}$$

والمطلوب ما يلي:

أ- بين إنَّ كان $X$ مستقل عن $Y$ أم لا؟

ب- احسب التباين لكل من المتغيِّرين العشوائيين $X$ و$Y$ وفقاً لتعريف التباين.

ج- عيِّن الدالَّة المولِّدة لعزوم المتغيِّر العشوائي $Y$، ومن ثمَّ استخدمها من أجل حساب التباين لهذا المتغيِّر العشوائي.

د- عيِّن الدالَّة المميِّزة للمتغيِّر العشوائي $X$، ومن ثمَّ استخدمها من أجل حساب التباين لهذا المتغيِّر العشوائي.

هـ- احسب قيمة مُعامل الارتباط الخطي بينهما وناقش قيمته.

و- إذا كانت $g$ دالّة حقيقية معرَّفة على $\mathbb{R}$ من خلال العلاقة $g(x) = (x-1)^2$ ، فعندئذ عيِّن الدالّة المولِّدة لعزوم المتغيِّر العشوائي $Z$ المُعطى بالعلاقة $Z = g(X) = (X-1)^2$ ، ومن ثمَّ استخدم هذه الدالّة في حساب $\mathbf{E}\,Z$ و $\mathbf{var}\,Z$ .

٣- اثبت أنَّ كلّ متغيِّر عشوائي تباينه معدوم سيكون خاضعاً للتوزيع الوحيد النقطة.

٤- ليكن $X$ متغيِّراً عشوائياً خاضعاً للتوزيع الطبيعي $\mathbf{N}(\mu , \sigma)$ والمطلوب ما يلي:

أ- حساب التباين لهذا المتغيِّر العشوائي.

ب- تعيين الدالّة المولِّدة لعزومه.

٥- عيِّن العزوم الابتدائية حتى المرتبة الرابعة للمتغيِّرات العشوائية الآتية مستخدماً دوالّها المميِّزة:

أ- $X$ متغيِّر عشوائي خاضع للتوزيع الوحيد النقطة عند القيمة $c = 0$ .

ب- $X$ متغيِّر عشوائي خاضع للتوزيع الحدّاني بمَعْلمتين $n = 50$ و $p = 0.25$ .

ج- $X$ متغيِّر عشوائي خاضع للتوزيع الأسّي بمَعْلمة $\lambda = 0.5$ .

٦- ليكن $X$ متغيِّراً عشوائياً خاضعاً للتوزيع المنتظم المتقطِّع بمَعْلمة $n \in \mathbb{N}$ ، والمطلوب تعيين الدالّة المولِّدة لعزومه وكذلك الدالّة المميِّزة له، ومن ثمَّ استخدمهما في حساب التوقع الرياضياتي والتباين لهذا المتغيِّر العشوائي.

٧- ليكن $X \sim \mathbf{G}(p)$ متغيِّراً عشوائياً، فعندئذ عيِّن الدالّة المولِّدة لعزومه، ومن ثمَّ استخدمها في حساب التوقع الرياضياتي والتباين لهذا المتغيِّر العشوائي.

٨- لنأخذ $X \sim \mathbf{EX}(\lambda)$ متغيِّراً عشوائياً، فعندئذ عيِّن دالّته المولِّدة الاحتمالية.

٩- استخدم خصائص الدالّة المولِّدة الاحتمالية لتوليد العزوم الابتدائية من المرتبة الأولى والثانية للمتغيِّرات العشوائية الآتية:

أ- المتغيِّر العشوائي $X \sim \mathbf{P}(\lambda)$ .

ب- المتغيِّر العشوائي $X \sim \mathbf{U}([a,b])$ .

ج- المتغيِّر العشوائي $X \sim \mathbf{EX}(\lambda)$ .

١٠- بفرض أنَّ $X$ و $Y$ متغيِّران عشوائيان يملكان توقُّعاً رياضياتياً مع $P(Y \neq 0) = 1$ و $\mathbf{E}Y \neq 0$ ، فعندئذ بيِّن من خلال مثال مناسب (غير الذي قدِّم في هذا الفصل) أنَّ $\mathbf{E}\dfrac{X}{Y} \neq \dfrac{\mathbf{E}X}{\mathbf{E}Y}$ .

١١- ليكن $X$ متغيِّراً عشوائياً مستمراً بدالّة كثافة احتمالية:

$$f_X(x) = \frac{1}{2}\mathbf{e}^{-|x|} \qquad ; \forall x \in \mathbb{R}$$

والمطلوب تعيين الدالّة المميِّزة $\varphi_X(t)$ لهذا المتغيِّر العشوائي.

١٢– ليكن $X$ متغيِّراً عشوائياً بدالَّة مميِّزة $\varphi_X(t) = \dfrac{1}{1+t^2}$ ، والمطلوب تعيين العزوم حتى المرتبة الثالثة لهذا المتغيِّر العشوائي.

١٣– بفرض أنَّ $X_1$ ، $X_2$ ، ...، و $X_n$ متغيِّرات عشوائية مستقلّة فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ ، فعندئذ أثبت صحة العلاقة الآتية:

$$\varphi_{X_1+X_2+\ldots+X_n}(t) = \prod_{k=1}^{n} \varphi_{X_k}(t) \qquad ;\forall\, t \in \mathbb{R}$$

١٤– تحت الفرضيات المقدَّمة في المبرهنة (٥ , ٢ , ٣ , ٧) اثبت صحة العلاقة الآتية:

$$\mathbf{G}_Y(s) = \mathbf{G}_\varsigma\left(\mathbf{G}_X(s)\right) \qquad ;\forall s \in \mathbb{R}$$

١٥– ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$ بدالَّة مولِّدة احتمالية $\mathbf{G}_X$ ، فعندئذ أثبت صحة العلاقة الآتية:

$$\mathbf{E}\left[X \cdot (X-1) \cdot \ldots \cdot (X-k+1)\right] = \mathbf{G}_X^{(k)}(1)$$

علماً أنَّ $\mathbf{G}_X^{(k)}(1)$ هي كتابة مختصرة للصيغة $\lim\limits_{s \uparrow 1} \mathbf{G}_X^{(k)}(s)$ .

١٦– ليكن $X \sim \mathbf{CO}(\mu\, ,\, \lambda)$ متغيِّراً عشوائياً، والمطلوب تعيين دالَّته المميِّزة $\varphi_X$ .

١٧– بالعودة إلى المثال / ١/ من (٤ , ٢ , ٥ , ٥) حيث لدينا $\mathbb{X}_2 = (X, Y)$ متَّجهاً عشوائياً متقطِّعاً بدالَّة احتمالية كتلية مُعطاة بالعلاقة الآتية:

$$p_{ij} = \begin{cases} \dfrac{7}{21} & for\ \ i = 1,2\ \&\ j = 1 \\[2mm] \dfrac{4}{21} & for\ \ i = 1\ \&\ j = 2 \\[2mm] \dfrac{3}{21} & for\ \ i = 2\ \&\ j = 2 \end{cases}$$

فهل المتغيِّر العشوائي $X$ مستقلّ عن $Y$ ؟

١٨– بالعودة إلى المثال / ٢/ من (٤ , ٢ , ٥ , ٥) حيث لدينا $\mathbb{X}_2 = (X, Y)$ متَّجهاً عشوائياً مستمرًّا بدالَّة كثافة احتمالية مُعطاة على النحو الآتي:

$$f_{\mathbb{X}_2}(x, y) = \begin{cases} 4x\,y & for\ \ (x, y) \in [0\,,1] \times [0\,,1] \\ 0 & otherwise \end{cases}$$

فهل المتغيِّران العشوائيان $X$ و $Y$ مستقلَّان عن بعضهما؟

١٩– بالعودة إلى التمرين / ١٤/ من تمارين الفصل الخامس حيث لدينا $\mathbb{X}_2 = (X, Y)$ متَّجهاً عشوائياً مستمرًّا بدالَّة كثافة احتمالية مُعطاة بالعلاقة الآتية:

$$f_{\mathbb{X}_2}(x, y) = \begin{cases} \alpha(x+y) & for\ \ (x, y) \in A \\ \\ 0 & OW. \end{cases}$$

فهل المتغيِّران العشوائيان $X$ و $Y$ مستقلَّان عن بعضهما؟

٢٠- بالعودة إلى التمرين / ١٥/ من تمارين الفصل الخامس حيث لدينا $\mathbb{X}_2 = (X, Y)$ متَّجهاً عشوائياً مستمرّاً بدالَّة كثافة احتمالية مُعطاة بالعلاقة الآتية:

$$f_{\mathbb{X}_2}(x, y) = \begin{cases} \mathbf{e}^{\alpha(x+y)} & for \quad (x, y) \in A \\ 0 & otherwise \end{cases}$$

فهل المتغيِّران العشوائيان $X$ و$Y$ مستقلَّان عن بعضهما؟

٢١- ليكن $X$ و$Y$ متغيِّرين عشوائيين بمجموعتي قيم $\mathbf{X} = \{x_i \,; i \in \mathbb{N}_k\}$ و $\mathbf{Y} = \{y_j \,; j \in \mathbb{N}_\ell\}$ على الترتيب، علماً أنَّ $k, \ell \in \mathbb{N}$ عددان محدودان ومثَّبتان. فإذا كان $X$ و$Y$ مستقلَّين ومتطابقي التوزيع، فعندئذ أثبت صحة العلاقة الآتية:

$$\mathbf{E}\left(X \mid X + Y\right) = \mathbf{E}\left(Y \mid X + Y\right) = \frac{X + Y}{2}$$

علماً أنَّ $P\left(X = x_i \,, Y = y_j\right) > 0$ من أجل كل $i \in \mathbb{N}_k$ وأيّ $j \in \mathbb{N}_\ell$.

# الفصل الثامن

## بعض نظريات النهاية المركزية

## SOME CENTRAL LIMIT THEOREMS

### تمهيد

لقد كانت دراساتنا في الفصول الأخيرة مركَّزة على عددمنته من المتغيِّرات العشوائية، ولكن الكثير من الدراسـات النظرية التي تسعى إلى تعميم نموذج ما من التجارب العشوائية من خلال أخذ القياسات، أو الملحوظات التي تتمخض عنها تكرارات كثيرة جداً لدراسة تجربة عشوائية ، أو ظاهرة عشوائية ما، كأن نفترض إمكانية إلقاء قطعة نقود معدنية لعددغير منته من المرّات، أو تكرار تجربة كيميائية لعددغير منته من المرّات، أو.... هنا نلاحظ أنَّ كلّ تكرار لتجربة يمثِّله متغيِّر عشوائيّ، فعلى سبيل المثال كلّ تكرار لتجربة إلقاء قطعة نقود معدنية يمثِّلها متغيِّر عشوائي برنوليّ بتوزيع معلوم مَعْلَمته $0 < p < 1$ تُحدَّد قيمته بحسب معلوماتنا عن المادة المصنَّعة منها القطعة المعدنية، فإن كانت المادة التي صنِّعت منها هذه القطعة المعدنية متجانسة، وعمليّة سبك النقد تتمّ بحيث يكون مركز ثقل القطعة المعدنية في مركزها تماماً، فعندئذ نفترض أنّها متوازنة، ويكون لدينا في هذه الحالة $p = \frac{1}{2}$ .

في الحقيقة إنَّ الدوافع وراء هذه الدراسات النظرية عديدة منها على سبيل المثال لا الحصر إيجاد حلول تقريبية لبعض مسائل الاحتمالات ذات الطابع التطبيقي، وذلك باستخدام توزيعات شهيرة معلومة التوزيع، وسهلة الحسـاب، وعلى أي حال، فهذه الدراسات الناتجة عن هذه التكرارات غير المنتهية (افتراضياً) تُنتِج لنا متتاليات لمتغيِّرات عشوائية بخصائص مُميَّزة (كالاستقلال مثلاً)، وقد تكون معلومة التوزيع في بعض الحالات عندها يَنصبّ الاهتمام على معرفة التوزيع الذي تؤول إليه هذه المتتالية. بمعنى أنَّه إذا كانت متتالية متغيِّرات عشوائية (من تلك التي ذُكِرت آنفاً) متقاربة فما هو توزيع المتغيِّر العشوائي الذي يمثِّل النهاية لهذه المتتالية؟

الآن، وبما أنَّ كلّ متغيِّر عشوائي هو تطبيق بقيم عشوائيَّة، فإنَّ دراسة توزيع المتغيِّر العشوائي الذي يمثِّل النهاية لمتتالية متغيِّرات عشوائية مُعطاة يتطلب تقديم مفاهيم مُحدَّدة لتقارب متتاليات المتغيِّرات العشوائيَّة أسوة بمفاهيم التقارب لمتتاليات الدوال. لذلك سنقوم فيما يلي بتقديم مبسط وموجز لبعض مفاهيم تقارب متتاليات المتغيِّرات العشوائية، وبعد ذلك الانتقال إلى البحث في سلوك المتغيِّر العشوائي الذي يمثِّل النهاية لمتتالية متغيِّرات عشوائية، وفي هذا الخصوص سنورد بعض المصطلحات ونذكِّر بالعديد من الحقائق المهمّة لتبسيط الدراسة (التي من المفترض أن يكون القارئ على اطلاع بمعظمها)، وكذلك سنقوم بتقديم بعض البراهين، ونتجاوز عن تلك البراهين الطويلة والمعقدَّة حيث يمكن لمن يودّ الاطلاع على تفاصيل هذه البراهين، أو يرغب في معرفة المزيد من المعلومات حول هذه المواضيع، الرجوع إلى المراجع ذات الصلة ومنها [10] و[14] و[24] و[28] و[39] و[46] و[53] في فهرس المراجع لهذا الكتاب.

سنبدأ دراستنا للموضوع المذكور أعلاه بتقديم بعض المفاهيم التي تتعلّق بتقارب متتاليات المتغيِّرات العشوائيَّة.

## (٨,١) تقارب متتاليات المتغيرات العشوائية
## Convergence of Sequences of Random Variables

في الواقع إنَّ دراسة التقارب لمتتاليات المتغيِّرات العشوائية تحتاج إلى تأسيس جيد في نظرية القياس (للمساعدة يمكن الاطلاع على الملحق B)، ولكن في دراستنا هذه سنوجز بعض الأفكار الضرورية لمسار دراستنا فقط.

### (٨,١,١) أشهر أنواع التقارب لمتتاليات المتغيِّرات العشوائيَّة

نقدِّم فيما يلي بعضاً من أنواع التقارب لمتتاليات المتغيِّرات العشوائيَّة ونبدأها بالتعريف الآتي.

**(٨,١,١,١) تعريف (التقارب بالتوزيع Converges in Distribution)**

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$. عندئذ يُقال عن المتتالية $(X_n)_{n\in\mathbb{N}}$ إنَّها **تتقارب بالتوزيع** من متغيِّر عشوائي $X$ فوق $[\Omega,\mathscr{A},P]$ (ويرمز لذلك بـ $X_n \xrightarrow{d} X$) إذا كان من أجل أيَّة دالَّة حقيقية $f$ مستمرَّة ومحدودة على $\mathbb{R}$ لدينا العلاقة الآتية مُحقَّقة:

$$\lim_{n\to\infty} \mathbf{E}\left[f\left(X_n\right)\right]=\mathbf{E}\left[f\left(X\right)\right] \qquad [8,1]$$

**(٨,١,١,٢) تعريف (التقارب بالاحتمال Converges in probability)**

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$. عندئذ يُقال إنَّ المتتالية $(X_n)_{n\in\mathbb{N}}$ **تتقارب بالاحتمال** من متغيِّر عشوائي $X$ فوق $[\Omega,\mathscr{A},P]$ (ويُرمز لذلك بـ $X_n \xrightarrow{P} X$) إذا كان من أجل كل $\varepsilon>0$ العلاقة الآتية مُحقَّقة:

$$\lim_{n\to\infty} P\left(\left|X_n-X\right|>\varepsilon\right)=0 \qquad [8,2]$$

**(٨,١,١,٣) تعريف (التقارب باحتمال يساوي الواحد Converges with probability equal to one)**

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$. عندئذ يُقال إنَّ المتتالية $(X_n)_{n\in\mathbb{N}}$ **تتقارب** من متغيِّر عشوائي $X$ فوق $[\Omega,\mathscr{A},P]$ **باحتمال يساوي الواحد** (ويرمز لذلك بـ $X_n \xrightarrow{a.s.} X$) إذا تحقَّقت العلاقة الآتية:

$$P\left(\lim_{n\to\infty} X_n=X\right)=1 \qquad [8,3]$$

وتجدر الإشارة هنا إلى أنَّ متتالية المتغيِّرات العشوائية المتقاربة باحتمال يساوي الواحد يقال عنها إنَّها **تتقارب تقريباً أكيدًا (أو شبه أكيد)** Converges Almost Sure أيضاً.

**(٨,١,١,٤) تعريف (التقارب بالمتوسط $p$ Converges in $p-$mean)**

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$. عندئذ يُقال إنَّ المتتالية $(X_n)_{n\in\mathbb{N}}$ **تتقارب بالمتوسط $p$** من متغيِّر عشوائي $X$ فوق $[\Omega,\mathscr{A},P]$ إذا تحقَّقت العلاقة الآتية:

$$\lim_{n\to\infty} \mathbf{E}|X_n-X|^p=0 \qquad [8,4\text{-a}]$$

علماً أنَّ $0<p<\infty$، ويُرمز لهذا التقارب بالشكل $X_n \xrightarrow{i.p.m.} X$، وفي الحالة الخاصة عندما يكون لدينا $p=2$ فإنَّ هذا النوع من التقارب يُدعى **التقارب بالمتوسط التربيعي** Convergence in Square Mean، وحينئذ يُرمز له بـ $X_n \xrightarrow{i.s.m.} X$، أو يكتب على النحو الآتي:

$$X=\underset{n}{l.i.m.}\, X_n \qquad [8,4\text{-b}]$$

### (٨، ١، ٢) تقارب متتاليات المتغيِّرات العشوائية

نقدِّم فيما يلي بعض المبرهنات المهمَّة (سنقبلها جميعاً دون برهان) التي توضِّح لنا حقائق عديدة تتعلَّق بهذه الأنواع من التقاربات بالإضافة إلى إظهار بعض الاقتضاءات الممكنة بين هذه الأنواع من التقاربات لمتتاليات المتغيِّرات العشـوائية.

#### (٨، ١، ٢، ١) مبرهنة

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$. عندئذ:

١- تكون هذه المتتالية متقاربة من متغيِّر عشوائي $X$ فوق $[\Omega,\mathscr{A},P]$ باحتمال يساوي الواحد **إذا وفقط إذا** كان من أجل كل $\varepsilon>0$ لدينا ما يلي مُحقَّقاً:

$$\lim_{n\to\infty} P\left(\sup_{k\geq n}|X_k-X|\geq\varepsilon\right)=0 \qquad [8,5]$$

٢- تكون هذه المتتالية **أساسية** Essentially Sequence (أو متتالية كوشي Cauchy Sequence) باحتمال يساوي الواحد إذا كان من أجل كل $\varepsilon>0$ لدينا ما يلي مُحقَّقاً:

$$\lim_{n\to\infty} P\left(\sup_{k\geq n\,,\,\ell\geq n}|X_k-X_\ell|\geq\varepsilon\right)=0 \qquad [8,6\text{-a}]$$

وهي تكافئ العلاقة الآتية أيضاً:

$$\lim_{n\to\infty} P\left(\sup_{k\geq 1}|X_{n+k}-X_n|\geq\varepsilon\right)=0 \qquad [8,6\text{-b}]$$

#### (٨، ١، ٢، ٢) نتيجة

بما أنَّ:

$$\left\{\omega\in\Omega\,;\sup_{k\geq n}|X_k(\omega)-X(\omega)|\geq\varepsilon\right\}=\left\{\omega\in\Omega\,;\bigcup_{k\geq n}|X_k(\omega)-X(\omega)|\geq\varepsilon\right\}$$

فإنَّه سيكون لدينا:

$$P\left(\sup_{k\geq n}|X_k-X|\geq\varepsilon\right)=P\left(\bigcup_{k\geq n}|X_k-X|\geq\varepsilon\right)\leq\sum_{k\geq n}P\left(|X_k-X|\geq\varepsilon\right)$$

ومن ثمَّ إذا كانت متتالية المتغيِّرات العشوائيَّة $(X_n)_{n\in\mathbb{N}}$ متقاربة من متغيِّر عشوائي $X$ فوق $[\Omega,\mathscr{A},P]$ باحتمال يساوي الواحد، فعندئذ سيكون من أجل كل $\varepsilon>0$ لدينا العلاقة الآتية محقَّقة:

$$\sum_{n=1}^{\infty}P\left(|X_n-X|\geq\varepsilon\right)<+\infty \qquad [8,7]$$

#### (٨، ١، ٢، ٣) مبرهنة

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$. عندئذ:

١- إذا كانت هذه المتتالية متقاربة من متغيِّر عشوائي $X$ بالاحتمال فإنَّها ستكون متقاربة من $X$ بالتوزيع أيضاً.

٢- إذا كانت هذه المتتالية متقاربة من $X$ باحتمال يساوي الواحد فإنَّها ستكون متقاربة من $X$ بالاحتمال أيضاً.

٣- إذا كانت هذه المتتالية متقاربة من $X$ بالمتوسط $p$ مع $0<p<\infty$ فإنَّها ستكون متقاربة من $X$ بالاحتمال أيضاً.

بناءً على نتائج هذه المبرهنة يمكننا أنَّ نعرض الاقتضاءات الآتية:

$$X_n \xrightarrow{a.s.} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$$

$$\Downarrow ? \quad \Uparrow ?$$

انتبه: إنَّ عكس هذه الاقتضاءات غير صحيح في الحالة العامّة.

$$X_n \xrightarrow{i.p.m} X \Rightarrow X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$$

وهكذا نجد أنّه في الحالة العامّة إذا كانت $X_n \xrightarrow{a.s.} X$ فإنَّ ذلك **لا** يقتضي أن تكون $X_n \xrightarrow{i.p.m} X$، وكذلك إذا كانت $X_n \xrightarrow{i.p.m} X$ فإنَّ ذلك **لا** يقتضي أن تكون $X_n \xrightarrow{a.s.} X$.

## (٨,٢) قانونا الأعداد الكبيرة

**Laws of Large Numbers**

من المواضيع المهمّة التي تربط نظرية الاحتمالات بالإحصاء الرياضياتي ما يُعرف باسم **قوانين الأعداد الكبيرة** التي تقدِّمها لنا التعاريف الآتية.

### (٨,٢,١) تعريف قانوني الأعداد الكبيرة (الضعيف و القوي)

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائيَّة فوق فضاء احتمالي $[\Omega, \mathscr{A}, P]$، ولنضع من أجل كل $n \in \mathbb{N}$ ما يلي:

$$S_n := X_1 + X_2 + \dots + X_n \qquad [8,8]$$

فعندئذ:

١- يُقال إنَّ المتتالية $(X_n)_{n\in\mathbb{N}}$ تخضع **للقانون الضعيف للأعداد الكبيرة** Weak Law of Large Numbers إذا وجد ثابت حقيقي $\mu$ بحيث يكون:

$$\frac{S_n}{n} \xrightarrow{P} \mu \qquad [8,9]$$

٢- يُقال إنَّ المتتالية $(X_n)_{n\in\mathbb{N}}$ تخضع **للقانون القوي للأعداد الكبيرة** Strong Law of Large Numbers إذا وجد ثابت حقيقي $\mu$ بحيث يكون:

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu \qquad [8,10]$$

#### (٨,٢,١,١) ملاحظات

١- من الملاحظ أنَّ قيمة $\frac{S_n}{n}$ تمثِّل القيمة الوسطى (أو المتوسط) للمتغيِّرات العشوائية $X_1$، $X_2$، ...، و $X_n$، ولذلك يُشار عادةً إلى $\frac{S_n}{n}$ بالرمز $\overline{X}_n$، أي إنَّه لدينا:

$$\overline{X}_n := \frac{S_n}{n} \quad ; \forall\ n \in \mathbb{N} \qquad [8,11\text{-}a]$$

٢- من أجل الحالة الخاصّة التي يكون فيها للمتغيِّرات العشوائية $X_1$، $X_2$، ...، و $X_n$ توزيع برنولي، فإنَّ $\frac{S_n}{n}$ يمثِّل التكرار النسبي لتَحقُّق وقوع الحادث الذي قيد الدراسة وذلك عند تكرار التجربة البرنوليّة تحت الشروط نفسها لـ $n$ مرَّة متتالية، ولذلك يُستخدم الرمز $\overline{P}_n$ للدلالة على $\frac{S_n}{n}$ لتمييز معناه عن $\overline{X}_n$، ولكي نتجنب الالتباس الذي يمكن أن ينشأ عن استخدام $\overline{P}_n$ (لتشابهه مع رمز الدالَّة الاحتماليّة $P$) فإنَّنا سنستخدم الرمز $\overline{\pi}_n$ بدلاً عن $\overline{P}_n$ للدلالة على $\frac{S_n}{n}$. أي إنَّنا سنكتب من أجل هذه الحالة الخاصّة:

$$\overline{\pi}_n := \frac{S_n}{n} \qquad ; \forall\, n \in \mathbb{N}$$ [8,11-b]

وإذا كانت هذه المتغيِّرات العشوائيَّة البرنوليّة مستقلّة ومتطابقة في التوزيع بمَعْلَمة مشترك $p$ (يمثّل احتمال النجاح في التجربة البرنوليّة الواحدة) فعندئذ يمثِّل المقدار $\frac{S_n}{n}$ عدد النجاحات خلال الـ $n$ تكرار.

٣- بما أنَّ التقارب باحتمال يساوي الواحد يقتضي التقارب بالاحتمال، فإنَّه إذا حَقَّقت متتالية متغيِّرات عشوائيَّة $(X_n)_{n\in\mathbb{N}}$ قانون الأعداد الكبيرة القوي فإنَّها ستكون مُحقِّقة لقانون الأعداد الكبيرة الضعيف حتماً، ولذلك يُركِّز المرء عادةً على الحالات التي تكون فيها $(X_n)_{n\in\mathbb{N}}$ مُحقِّقة لقانون الأعداد الكبيرة القوي.

المبرهنة الآتية التي سنبرهن على إحدى نتيجتيها فقط (لمن يود الاطلاع على برهان النتيجة الأولى منها يمكنه الرجوع إلى بعض المراجع التي ذكرت في بداية هذا الفصل) تقدِّم لنا ثمرة من ثمار قوانين الأعداد الكبيرة.

### (٨, ٢, ٢) مبرهنة بخصوص تقارب المتتالية $\left(\frac{S_n}{n}\right)_{n\in\mathbb{N}}$

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائيَّة فوق فضاء احتمالي $[\,\Omega, \mathscr{A}, P\,]$، فإذا كانت عناصر هذه المتتالية مستقلّة ومتطابقة في التوزيع وتملك عزماً ابتدائياً من المرتبة الثانية، فعندئذ يكون لدينا ما يلي مُحقَّقاً:

$$\frac{S_n}{n} \xrightarrow{a.s.} \mathbf{E}X_1$$ [8,12-a]

$$\frac{S_n}{n} \xrightarrow{i.s.m.} \mathbf{E}X_1$$ [8,12-b]

**برهان النتيجة [8,12,b]:**

بما أنَّ عناصر المتتالية $(X_n)_{n\in\mathbb{N}}$ مستقلّة، ومتطابقة التوزيع، وتملك عزماً ابتدائياً من المرتبة الثانية، فإنَّه يُمكننا أن نكتب ما يلي:

$$\mathbf{E}\left(\frac{S_n}{n} - \mathbf{E}X_1\right)^2 = \mathbf{E}\left(\frac{S_n}{n} - \frac{n\ \mathbf{E}X_1}{n}\right)^2 = \mathbf{E}\left(\frac{S_n}{n} - \frac{1}{n}\sum_{k=1}^{n}\mathbf{E}X_1\right)^2 = \mathbf{E}\left(\frac{S_n}{n} - \frac{\mathbf{E}X_1 + \mathbf{E}X_2 + \ldots + \mathbf{E}X_n}{n}\right)^2$$
$$= \mathbf{E}\left(\frac{S_n}{n} - \frac{\mathbf{E}(X_1 + X_2 + \ldots + X_n)}{n}\right)^2 = \mathbf{E}\left[\frac{1}{n^2}\left(S_n - \mathbf{E}S_n\right)^2\right] = \frac{1}{n^2}\ \mathbf{var}\, S_n$$

وباستخدام خاصّيّة الاستقلال للمتغيِّرات العشوائيَّة في المتتالية $(X_n)_{n\in\mathbb{N}}$ يصبح لدينا ما يلي:

$$\mathbf{E}\left(\frac{S_n}{n} - \mathbf{E}X_1\right)^2 = \frac{1}{n^2}\sum_{k=1}^{n}\mathbf{var}\, X_k = \frac{1}{n}\ \mathbf{var}\, X_1$$

ومن ثَمّ ينتج لدينا من كون $0 < \mathbf{var}\, X_1 < \infty$ ما يلي مُحقَّقاً:

$$\lim_{n\to\infty}\ \mathbf{E}\left(\frac{S_n}{n} - \mathbf{E}X_1\right)^2 = \lim_{n\to\infty}\ \frac{\mathbf{var}\, X_1}{n} = 0$$

وهذا يعني أنَّ $\frac{S_n}{n} \xrightarrow{i.s.m.} \mathbf{E}X_1$، وبذلك يتمّ البرهان.

**(٨,٢,٢,١) مثال**

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائيَّة فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ مستقلّةوبرنوليّة التوزيع، ولنفترض أنّه من أجل كل $n\in\mathbb{N}$ لدينا $P(X_n=1)=p_n$ و $P(X_n=0)=1-p_n$، علماً أنَّ $0<p_n<1$، فعندئذ بحسب الدراسة السابقة يكون لدينا $\frac{1}{n}\sum_{k=1}^{n}p_k \xrightarrow{P} \overline{\pi}_n$، وذلك لأنّه من أجل كل $n\in\mathbb{N}$ لدينا:

$$\mathbf{E}\,\overline{\pi}_n=\frac{1}{n}\sum_{k=1}^{n}\mathbf{E}X_k=\frac{1}{n}\sum_{k=1}^{n}p_k$$

وكذلك:

$$\mathbf{var}\,\overline{\pi}_n=\frac{1}{n^2}\mathbf{var}\,S_n=\frac{1}{n^2}\sum_{k=1}^{n}\mathbf{var}\,X_k=\frac{1}{n^2}\sum_{k=1}^{n}p_k(1-p_k)$$

وباستخدام متباينة تشبيشيف (انظر العلاقة [7,36-a] في المبرهنة (٧,٢,٥,١)) يكون لدينا من أجل كل $\varepsilon>0$ ما يلي مُحقَّقاً:

$$P\left(\left|\overline{\pi}_n-\mathbf{E}\overline{\pi}_n\right|\geq\varepsilon\right)\leq\frac{\sqrt{\mathbf{var}\,\overline{\pi}_n}}{\sqrt{\varepsilon}}\leq\frac{1}{2\sqrt{\varepsilon\cdot n}}\xrightarrow[n\to\infty]{}0$$

وهذا يعني أنَّ المتتالية $(\overline{\pi}_n)_{n\in\mathbb{N}}$ تتقارب بالاحتمال من القيمة $\mathbf{E}(\overline{\pi}_n)$.

تجدر الإشارة هنا إلى أنّه إذا كانت متتالية المتغيِّرات العشوائيَّة بالمواصفات الواردة في متن هذا المثال تُحقِّق التقارب $\frac{S_n}{n}\xrightarrow{a.s.}\mathbf{E}X_1$، فعندئذ يُقال إنَّ هذه المتتالية تحقِّق **قانون بواسون للأعداد الكبيرة** Poisson's Law of Large Numbers، وفي الحالة الخاصّة عندما يكون $p_n=p$ من أجل كل $n\in\mathbb{N}$ فإنَّه يُقال عن هذه المتتالية إنّها تُحقِّق **قانون برنولي للأعداد الكبيرة** Bernoulli's Law of Large Numbers.

**(٨,٢,٢,٢) ملاحظة**

إذا كانت متتالية المتغيِّرات العشوائيَّة $(X_n)_{n\in\mathbb{N}}$ **لا** تملك عزوماً ابتدائيةً ولا من أية مرتبة، فإنَّ المبرهنة السابقة تصبح عديمة التطبيق، وفي مثل هذه الحالات يَلجأ المرء إلى تقنية الدوال المميّزة، وهذا ما ستظهره المبرهنة الآتية التي سنقبلها دون برهان أيضاً.

**(٨,٢,٦) مبرهنة بخصوص تقارب المتتالية $\left(\frac{S_n}{n}\right)_{n\in\mathbb{N}}$ من ثابت حقيقي $\mu$**

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائيَّة فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ مستقلّة ومتطابقة في التوزيع، فعندئذ من أجل $\mu$ ثابت يكون لدينا $\frac{S_n}{n}\xrightarrow{P}\mu$ **إذا وفقط إذا** تحقَّق أحد الشرطين الآتيين:

١- لدينا:

$$\text{a)}\quad\lim_{n\to\infty}n\cdot P(|X_1|>n)=0\qquad\&\qquad\text{b)}\quad\lim_{n\to\infty}\mathbf{E}\left(X_1\cdot\mathbf{I}_{\{|X_1|\leq n\}}\right)=\mu$$

٢- الدالَّة المميّزة $\varphi_{X_n}(t)$ قابلة للمفاضلة في النقطة $t=0$ من أجل كل $n\in\mathbb{N}$، ولدينا:

$$\varphi'_{X_n}(0)=\mathbf{i}\cdot\mu\qquad;\forall\,n\in\mathbb{N}$$

فنلاحظ هنا أنّه لم يُشْترط وجود العزوم لعناصر متتالية المتغيِّرات العشوائيَّة $(X_n)_{n\in\mathbb{N}}$ لكي يتحقَّق القانون الضعيف للأعداد الكبيرة.

## (٨,٣) نظريات النهاية المركزية

**Central Limit Theorems**

إنَّ نظريات النها ية المركزية تُعدّ من النظريات المهمّة جداً على صعيد الدراسات العشوائية التطبيقية والنظرية على حد سواء، وعلى وجه الخصوص في مجال الدراسات الإحصائية، وهذه النظريات تعمل تحت شروط متنوعة بحسب الغاية التي تتطلبها المسألة المدروسة، وذلك من أجل حساب احتمالات لحوادث تتعلّق بتجربة عشوائية ما أو استقراءٍ يتعلّق بمجتمع إحصائي ما لدى الدراسات الإحصائية. إنَّ تقنية هذه النظريات تقوم على تعيين دوال التوزيع لمتتالية دوال توزيع لأسرة عدودة من المتغيِّرات العشوائية المستقلّة. كذلك سنرى أنَّ دراسة نظريات النهاية المركزية مبنيّة على الصلة الوطيدة ما بين نهايات المتتاليات للدوال المميِّزة للمتغيِّرات العشوائية ونهايات المتتاليات لدوال التوزيع الاحتمالية للمتغيِّرات العشوائية، وفي هذا الصدد سنلاحظ أنَّ التوزيع الطبيعي سيلعب دوراً مميّزاً ومركزياً في هذا النوع من الدراسات أيضاً.

في الحقيقة من أهمّ الأسباب التي تدفعنا إلى استخدام نظريات النهاية المركزية هو أنّه عندما يصبح عدد المتغيِّرات العشوائية التي تُنمذج الظاهرة التي قيد الدراسة كبيراً بقدر كاف، وكان لدينا دالّة (قيوسة) $g$ بهذه المتغيِّرات العشوائية، فعندئذ لا يعود اهتمامنا مركَّزاً على أن تأخذ هذه الدالّة $g$ قيمة مُحدَّدة، وإنّما يتركّز اهتمامنا على أن يكون لهذه الدالّة $g$ قيمة في فترة $I$ من $\mathbb{R}$ .

سنقدِّم فيما يلي بعض نظريات النهاية المركزية التي سنبدأها بأبسط نموذج منها، ألا وهو نظرية موافير– لابلاس.

### (٨,٣,١) مسألة موافير– لابلاس Moivre – Laplace Problem:

إنَّ هذه المسألة تنطلق من حالة خاصّة جداً، وذلك عندما تكون متتالية المتغيِّرات العشوائية التي قيد الدراسة مستقلّة ومتطابقةً في التوزيع مع توزيع متغيِّر عشوائي $X$ خاضع لتوزيع برنولي، والنظرية الآتية توضِّح لنا ذلك.

#### (٨,٣,١,١) نظرية موافير– لابلاس Moivre – Laplace Theorem:

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائيَّة فوق فضاء احتمالي $[\,\Omega,\mathscr{A},P\,]$، ولنضع ما يلي:

$$Z_n=\frac{S_n-\mathbf{E}S_n}{\sqrt{\mathbf{var}\,S_n}}\qquad ;\ \forall\ n\in\mathbb{N}\qquad [8,13]$$

علماً أنَّ $S_n$ معطاة بالعلاقة [8,8] ، كما نلاحظ هنا أنَّ $\mathbf{E}Z_n=0$ و $\mathbf{var}\,Z_n=1$ من أجل كل $n\in\mathbb{N}$ . الان، إذا كانت المتغيِّرات العشوائيَّة $X_1$ ، $X_2$ ،...و $X_n$ و... مستقلّة وبرنوليّة بمَعْلَمة مشتركة $0<p<1$ (أي إنّها متطابقةً في التوزيع مع توزيع متغيِّر عشوائي $X$ خاضع لتوزيع برنولي بمَعْلَمة $0<p<1$)، فعندئذ ستكون العلاقة الآتية مُحقَّقة من أجل أيّة قيمة $z\in\mathbb{R}$ :

$$\lim_{n\to\infty}F_{Z_n}(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}\mathbf{e}^{-\frac{t^2}{2}}\,dt\ =\ \Phi(z)\qquad [8,14]$$

البرهان: لنفترض أنَّ $\varphi_{S_n}$ هي الدالّة المميِّزة للمتغيِّر العشوائي $S_n$ من أجل كل $n\in\mathbb{N}$ ، فعندئذ بسبب أنّه للمتغيِّر العشوائي $S_n$ توزيعًا حدَّانيًا بمَعْلمتين $n$ و$p$ ، فإنّه سينتج من استقلال المتغيِّرات العشوائية $X_1$ ، $X_2$ ،...و $X_n$ و... العلاقة الآتية:

$$\varphi_{S_n}(t)=\left[\,1-p\cdot\left(1-\mathbf{e}^{\mathbf{i}t}\,\right)\,\right]^n\qquad ;\ \forall t\in\mathbb{R}$$

ومن ثمّ يكون:

$$\varphi_{Z_n}(t)\ =\exp\left[\frac{-\mathbf{i}t\cdot\mathbf{E}S_n}{\sqrt{\mathbf{var}\,S_n}}\right]\cdot\varphi_{S_n}\left(\frac{t}{\sqrt{\mathbf{var}\,S_n}}\right)$$

ولكن لدينا:

$$\mathbf{exp}\left[\frac{-\mathbf{i}t\cdot\mathbf{E}S_n}{\sqrt{\mathbf{var}S_n}}\right]\cdot\varphi_{S_n}\left(\frac{t}{\sqrt{\mathbf{var}S_n}}\right)=\left[(1-p)\cdot\mathbf{exp}\left(\frac{-\mathbf{i}pt}{\sqrt{np(1-p)}}\right)+p\cdot\mathbf{exp}\left(\frac{\mathbf{i}t(1-p)}{\sqrt{np(1-p)}}\right)\right]^n$$

ومن نشر تايلور للدالَّة $\mathbf{e}^{\mathbf{i}t}$ يمكننا أنَّ نكتب $\mathbf{e}^{\mathbf{i}t}=\sum_{k=0}^{n}\frac{(\mathbf{i}t)^k}{k\,!}+\theta(t^n)$ علماً أنَّ $\theta(t^n)$ مقدار لا متناهي في الصغر من الدرجة $n$ بـ $t$، ومن ثمَّ يمكننا من أجل كل $t\in\mathbb{R}$ أنَّ نكتب الآتي:

$$p\cdot\mathbf{exp}\left(\frac{\mathbf{i}t(1-p)}{\sqrt{np(1-p)}}\right)=p+\mathbf{i}t\sqrt{\frac{p(1-p)}{n}}-\frac{(1-p)t^2}{2n}+\theta\left(\frac{t^2}{n}\right)$$

$$(1-p)\cdot\mathbf{exp}\left(\frac{-\mathbf{i}tp}{\sqrt{np(1-p)}}\right)=(1-p)-\mathbf{i}t\sqrt{\frac{p(1-p)}{n}}-\frac{pt^2}{2n}+\theta\left(\frac{t^2}{n}\right)$$

علماً أنَّ $\lim\limits_{n\to\infty} n\cdot\theta\left(\frac{1}{n}t^2\right)=0$، فلو قمنا الآن بالتعويض في صيغة $\varphi_{Z_n}$ مع مراعاة أنَّ $p+(1-p)=1$ فإنَّنا سنحصل على العلاقة الآتية:

$$\varphi_{Z_n}(t)=\left[1-\frac{t^2}{2n}+\theta\left(\frac{t^2}{n}\right)\right]^n \qquad ;\forall t\in\mathbb{R}$$

وبأخذ لوغاريتم الطرفين للعلاقة السابقة يصبح لدينا:

$$\mathbf{ln}\,\varphi_{Z_n}(t)=n\cdot\mathbf{ln}\left[1-\frac{t^2}{2n}+\theta\left(\frac{t^2}{n}\right)\right] \qquad ;\forall t\in\mathbb{R}$$

وبوضع $\alpha:=\frac{-t^2}{2n}+\theta\left(\frac{1}{n}t^2\right)$، وبسبب أنَّ $|\alpha|<1$، ومن ثمَّ نشر الدالَّة $\mathbf{ln}(1+\alpha)$، فإنَّه من أجل أيّ $t\in\mathbb{R}$ مثبَّتة و $n$ كبيرة بقدرٍ كاف، ينتج لدينا ما يلي:

$$\mathbf{ln}\,\varphi_{Z_n}(t)=n\cdot\mathbf{ln}\,(1+\alpha)=n\cdot\alpha=\frac{-t^2}{2}+n\cdot\theta\left(\frac{t^2}{n}\right)$$

وبأخذ نهاية الطرفين للعلاقة السابقة عندما $n\to\infty$ نحصل على العلاقة الآتية:

$$\lim_{n\to\infty}\mathbf{ln}\,\varphi_{Z_n}(t)=\frac{t^2}{2} \qquad ;\forall t\in\mathbb{R}$$

وهذا يعني أنَّ:

$$\lim_{n\to\infty}\varphi_{Z_n}(t)=\mathbf{e}^{-\frac{t^2}{2}} \qquad ;\forall t\in\mathbb{R}$$

وهذه العلاقةالأخيرة ما هي إلاَّ الدالَّة المميِّزة لمتغيِّر عشوائي خاضع للتوزيع الطبيعي المعيّاري، ومنه بحسب مبرهنة الاستمرار لـ ليفي / ٣-٣-٤-٧/ فإنَّه سينتج لدينا العلاقة الاتية التي تنهي برهان المبرهنة:

$$\lim_{n\to\infty}F_{Z_n}(z)=\Phi(z)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z}\mathbf{e}^{-\frac{t^2}{2}}\,dt \qquad ;\forall t\in\mathbb{R}$$

### (٨,٣,١,٢) نتائج

١- إذا كانت $n$ كبيرة بقدر كاف، فعندئذ يمكننا باستخدام النظرية السابقة أن نكتب الآتي من أجل أيّة قيمتين $z_1 , z_2$ و $\mathbb{R}$ مع $z_2 > z_1$:

$$P\left(z_1 < Z_n < z_2\right) \approx \frac{1}{\sqrt{2\pi}} \int_{z_1}^{z_2} \mathbf{e}^{-\frac{t^2}{2}}\, dt = \Phi(z_2) - \Phi(z_1) \qquad [8,15]$$

التي نستنتج منها أنَّ:

$$P\left(z_1\sqrt{n\,p\,(1-p)} + n\,p < S_n < z_2\sqrt{n\,p\,(1-p)} + n\,p\right) \approx \Phi(z_2) - \Phi(z_1)$$

فإذا افترضنا أنَّ:

$$s_1 = z_1\sqrt{n\,p\,(1-p)} + n\,p \quad \& \quad s_2 = z_2\sqrt{n\,p\,(1-p)} + n\,p \qquad (*)$$

فعندئذ سيصبح لدينا:

$$P(s_1 < S_n < s_2) \approx \Phi\left(\frac{s_2 - n\,p}{\sqrt{n\,p\,(1-p)}}\right) - \Phi\left(\frac{s_1 - n\,p}{\sqrt{n\,p\,(1-p)}}\right) \qquad [8,16]$$

ويُقال في هذه الحالة إنَّ $S_n$ **مقارب للتوزيع الطبيعي** Asymptotically Normal Distributed بمتوسط $n \cdot p$ وانحراف معياري $\sqrt{n \cdot p \cdot (1-p)}$، وسنكتب ذلك اصطلاحاً بالشكل $S_n \overset{\infty}{\sim} \mathbf{N}\left(n\,p\,,\sqrt{n\,p\,(1-p)}\right)$.

٢- بسبب أنَّ $S_n \sim \mathbf{B}(n,p)$ فإنَّه سيكون لدينا:

$$\mathbf{E}\,\overline{\pi}_n = p \quad \& \quad \mathbf{var}\,\overline{\pi}_n = \frac{p\,(1-p)}{n} \qquad [8,17]$$

ومن ثَمَّ سيصبح المتغيِّر العشوائي الآتي معيّارياً:

$$Z_n := \frac{\overline{\pi}_n - \mathbf{E}\,\overline{\pi}_n}{\sqrt{\mathbf{var}\,\overline{\pi}_n}} \qquad [8,18]$$

ومن ثَمّ ستُحقِّق متتالية دوال التوزيع $\left(F_{Z_n}\right)_{n\in\mathbb{N}}$ نظرية **موافير- لابلاس**، ومنه ينتج لدينا أنَّ:

$$\lim_{n\to\infty} F_{Z_n}(z) = \Phi(z) \qquad ; \forall z \in \mathbb{R}$$

وبناء على ما سبق فإنَّه من أجل أيّة قيمتين $z_1$ و $z_2$ من $\mathbb{R}$ مع $z_2 > z_1$ يمكننا أنَّ نكتب الآتي:

$$P\left(z_1 < Z_n < z_2\right) \approx \Phi(z_2) - \Phi(z_1)$$

ولكن لدينا:

$$P(z_1 < Z_n < z_2) = P\left(z_1 < \frac{\overline{\pi}_n - p}{\sqrt{\frac{p\,(1-p)}{n}}} < z_2\right) = P\left(z_1\sqrt{\frac{p\,(1-p)}{n}} + p < \overline{\pi}_n < z_2\sqrt{\frac{p\,(1-p)}{n}} + p\right)$$

فلو وضعنا:

$$\beta_1 = z_1\sqrt{\frac{p\,(1-p)}{n}} + p \quad \& \quad \beta_2 = z_2\sqrt{\frac{p\,(1-p)}{n}} + p$$

فإنَّه يصبح لدينا ما يلي:

$$P\left(\beta_1 < \overline{\pi}_n < \beta_2\right) \approx \Phi\left(\frac{\beta_2 - p}{\sqrt{\frac{p(1-p)}{n}}}\right) - \Phi\left(\frac{\beta_1 - p}{\sqrt{\frac{p(1-p)}{n}}}\right) \qquad [8,19]$$

وفي هذه الحالة يُقال عن المتغيِّر العشوائي $\overline{\pi}_n$ إنَّه **مقارب للتوزيع الطبيعي** بمتوسط $p$ وانحراف معياري $\sqrt{\frac{p(1-p)}{n}}$، وسنكتب ذلك اصطلاحاً على النحو الآتي:

$$\overline{\pi}_n \overset{\infty}{\sim} \mathbf{N}\left(p \,,\, \sqrt{\frac{p(1-p)}{n}}\right)$$

**(٣,١,٣,٨) أمثلة**

١- لدى تنفيذ استطلاع للرأي حول جودة منتج صناعي طرح سؤال يحتمل إحدى إجابتين فقط (جيد أو رديء) على 1000 شخص وبشكل مستقل كل منهم عن الآخر، فإذا افترضنا أنَّ أيّ الإجابتين لها الاحتمال نفسه، فما هو احتمال أن يكون عدد الذين أجابوا بعبارة **"جيد"** ما بين 525 و550 شخصاً؟

الحـل: من أجل ذلك لنفترض أنَّ $X_n$ متغيِّر عشوائي راصد لإجابة الشخص ذي الرقم $n$ مع $n \in \mathbb{N}_{1000}$ فعندئذ نجد أنَّ هذه المتغيِّرات العشوائيَّة هي متغيِّرات عشوائيَّة برنوليّة ومستقلّة بمَعْلَمة مشترك $p = \frac{1}{2}$ (ومتطابقة في التوزيع أيضاً)، فلو افترضنا أنَّ عبارة جيد ورديء يقابلها القيمة 1 و 0 على الترتيب، فإنَّه سيكون لدينا:

$$P(X_n = 1) = P(X_n = 0) = \frac{1}{2}$$

فلو أخذنا $S_n$ من أجل كل $n \in \mathbb{N}$ متغيِّر عشوائي مُعطى وفقاً للعلاقة [8,8]، فعندئذ سيكون لـ $S_n$ توزيع حدَّاني $\mathbf{B}\left(1000 \,,\, \frac{1}{2}\right)$، ومن ثَمّ ينتج لدينا أنَّ $\mathbf{E}S_n = 500$ و $\mathbf{var}\,S_n = 250$، ومن ثمَّ يكون الاحتمال المطلوب يساوي:

$$\begin{aligned} P(525 < S_n < 550) &= P\left(\frac{525-500}{\sqrt{250}} < \frac{S_n - 500}{\sqrt{250}} < \frac{550-500}{\sqrt{250}}\right) \\ &= P\left(1.58 < \overline{\pi}_n < 3.16\right) \approx \Phi(3.16) - \Phi(1.58) \approx 1 - 0.94298 = 0.05702 \end{aligned}$$

علماً أنَّ $\overline{\pi}_n := \frac{S_n - 500}{\sqrt{250}}$، وبهذا ينتهي حل المثال.

٢- في بلد قامت إحدى المؤسسات الإعلامية باستطلاع للرأي حول موضوع البطالة، وقد شمل الاستطلاع 75 ألف شخص اختيروا عشوائياً، حيث أجاب 15 ألفاً منهم بـ **" نعم "** (أي إنَّه عاطل عن العمل)، ومن ثمَّ ادعت هذه المؤسسة أنَّ نسبة العاطلين عن العمل تُشكِّل 20% من التعداد العام للسكان. لكن الهيئة المسؤولة عن العمل في ذلك البلد شكَّكت في هذه النتيجة لاعتقادها أن الاستطلاع لم يشمل عدداً كافياً من الأشخاص. فطلبت من مكتب الإحصاء لديها أن يقوم بحساب عدد الأشخاص $n$ الذي يجب أن يشملهم الاستطلاع حتى تكون نسبة العاطلين عن العمل ما بين 19% و 21% باحتمال قدره 0.95 حيث ستقبل بأنَّ هذه القيمة للاحتمال تعطي مصداقية كافية بقبول العدد $n$ ومن ثمَّ مقارنته مع عدد الأشخاص الذين شملهم الاستطلاع لاتخاذ قرار القبول أو الرفض.

الحـل: للإجابة على هذا الطلب سنفترض $X_k$ متغيِّر عشوائي راصد لنتيجة جواب الشخص ذي الرقم $k$ مع $k \in \mathbb{N}_n$، فعندئذ نجد أنَّ هذه المتغيِّرات العشـوائيَّة هي متغيِّرات عشوائيَّة برنوليّة مستقلّة مع $P(X_n = 0) = 0.80$ و $P(X_n = 1) = 0.20$،

حيث لدينا هنا النجاح هو الحصول على شخص عاطل عن العمل، وآخذين في الحسبان أنَّ نسبة العاطلين عن العمل تشكل 20% من التعداد العام للسكان وفقاً للاستطلاع الذي تمّ، ومن ثمَّ يكون لهذه المتغيِّرات العشـوائيَّة مَعْلَمة مشترك $p=0.20$ (وهذه المتغيِّرات العشوائيَّة متطابقة في التوزيع أيضاً).

الآن لو أخذنا من أجل كل $n \in \mathbb{N}$ المتغيِّرين العشوائيين $S_n$ و $\overline{\pi}_n$ المُعطيين بالعلاقتين [8,8] و [8,11-b] على الترتيب، فعندئذ يُمثِّل المتغيِّر العشوائي $\overline{\pi}_n$ التكرار النسبي للعاطلين عن العمل، ومن ثمَّ يكون لدينا:

$$\mathbf{E}\,\overline{\pi}_n = \frac{1}{n}\mathbf{E}\,S_n = \frac{n\,p}{n} = p = 0.20$$

$$\mathbf{var}\,\overline{\pi}_n = \frac{1}{n^2}\mathbf{var}\,S_n = \frac{n\,p\,(1-p)}{n^2} = \frac{p\,(1-p)}{n} = \frac{0.16}{n}$$

ولكن من الفرض لدينا:

$$\begin{aligned} 0.95 = P\left(0.19 < \overline{\pi}_n < 0.21\right) &= P\left(\frac{0.19-0.20}{\sqrt{\dfrac{0.16}{n}}} < \frac{\overline{\pi}_n - 0.20}{\sqrt{\dfrac{0.16}{n}}} < \frac{0.21-0.20}{\sqrt{\dfrac{0.16}{n}}}\right) \\ &= P\left(-0.25\sqrt{n} < \frac{\overline{\pi}_n - 0.20}{\sqrt{\dfrac{0.16}{n}}} < 0.25\sqrt{n}\right) \approx \frac{1}{\sqrt{2\pi}}\int_{-0.25\sqrt{n}}^{0.25\sqrt{n}} \mathbf{e}^{-\frac{t^2}{2}}\,dt \end{aligned}$$

وبوضع $z = 0.25\sqrt{n}$ يصبح لدينا:

$$\frac{1}{\sqrt{2\pi}}\int_{-z}^{z} \mathbf{e}^{-\frac{t^2}{2}}\,dt = 2\frac{1}{\sqrt{2\pi}}\int_{0}^{z} \mathbf{e}^{-\frac{t^2}{2}}\,dt = 2\left(\Phi(z) - 0.5\right)$$

ومن ثمَّ يكون $\Phi(z) - 0.5 = 0.475$، وبالتالي ينتج أنَّ $\Phi(z) = 0.975$، وباستخدام جدول التوزيع الطبيعي المعياري نجد أنَّ قيمة $z$ تساوي 1.96، ومنه يكون لدينا $0.25\sqrt{n} = 1.96$ والتي نحصل منها على حجم العينة المفترض سحبها $n \approx 61.466$ ، ومن ثمَّ يجب أنَّ يكون عدد الأشخاص الذين يشملهم الاستطلاع لا يقل عن 61466 شخصاً، وهو أقل من العدد الذي تمَّ سبره في استطلاع الرأي، ومن ثمَّ فليس أمام الهيئة المسؤولة عن العمل من خيار إلاّ القبول بنتائج هذا الاستطلاع.

### (٨,٣,٢) مسألة ليندنبرغ – ليفي Lindenberg – Levy Problem

إنَّ هذه المسألة تنطلق من حالة أقل خصوصيّة من سابقتها وذلك عندما تكون متتالية المتغيِّرات العشوائية المستقلّة متطابقةً في التوزيع ولكن ليس بالضرورة أن تكون برنوليّة التوزيع، ومقابل ذلك التعميم ستظهر لدينا قيودًا إضافية توضِّحها لنا المبرهنة التي سنقدِّمها بعد التمهيد الآتي.

#### (٨,٣,٢,١) تمهيد

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية فوق فضاء احتمالي $[\,\Omega,\mathscr{A},P\,]$ مستقلّة ومتطابقة التوزيع مع توزيع متغيِّر عشوائي $X$، وتملك عزوماً ابتدائيةً من المرتبة الثانية، ولنفترض أنّه من أجل كل $n \in \mathbb{N}$ لدينا:

$$\mathbf{E}\,X_n = \mathbf{E}\,X = \mu \qquad \& \qquad \mathbf{var}\,X_n = \mathbf{var}\,X = \sigma^2$$

فلو أخذنا $S_n$ من أجل كل $n \in \mathbb{N}$ متغيِّر عشوائي مُعطى وفقاً للعلاقة [8,8] ، فعندئذ سيكون من أجل كل $n \in \mathbb{N}$ لدينا:

$$\mathbf{E}\,S_n = n \cdot \mu \qquad \& \qquad \mathbf{var}\,S_n = n \cdot \sigma^2$$

ولنقم باستعيار المتغيِّرات العشوائية $S_n$ باستخدام العلاقة [8,13] فيكون لدينا:

$$Z_n := \frac{S_n - \mathbf{E}\,S_n}{\sqrt{\mathbf{var}\,S_n}} = \frac{S_n - n \cdot \mu}{\sigma \cdot \sqrt{n}} \qquad ;\forall n \in \mathbb{N} \qquad [8,20]$$

توضِّح لنا النظرية الآتية أنَّ متتالية دوال التوزيع $(F_{Z_n})_{n\in\mathbb{N}}$ للمتغيِّرات العشوائية المعرَّفة بالعلاقة السابقة [8,20] ستكون متقاربة من دالَّة التوزيع الطبيعي المعياري $\Phi$ .

### (٢ ,٢ ,٣ ,٨) نظرية ليندنبرغ- ليفي Lindenberg–Levy Theorem

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega, \mathcal{A}, P]$ مستقلّة ومتطابقة التوزيع مع توزيع متغيِّر عشوائي $X$ ، وتملك عزوماً ابتدائية من المرتبة الثانية مع:

$$\mathbf{E}\,X_n = \mathbf{E}\,X = \mu \qquad \& \qquad \mathbf{var}\,X_n = \mathbf{var}\,X = \sigma^2 > 0 \qquad ;\forall n \in \mathbb{N}$$

ولتكن $(Z_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية عناصرها معرَّفة بالعلاقة [8,20] ، ولنفترض أنَّ $F_{Z_1}$ ، $F_{Z_2}$ ،...، $F_{Z_n}$ و... هي متتالية دوال التوزيع الاحتمالية للمتغيِّرات العشوائية $Z_1$ ، $Z_2$ ،...، $Z_n$ و... على الترتيب، فعندئذ ستكون العلاقة الآتية مُحقَّقة:

$$\lim_{n\to\infty} F_{Z_n}(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-\frac{t^2}{2}}\, dt = \Phi(z) \qquad ;\forall n \in \mathbb{N} \qquad [8,21]$$

البرهان: بحسب مُعطيات المسألة يمكننا أن نكتب من أجل كل $n \in \mathbb{N}$ ما يلي:

$$Z_n = \frac{S_n - \mathbf{E}\,S_n}{\sqrt{\mathbf{var}\,S_n}} = \frac{S_n - n \cdot \mu}{\sigma \cdot \sqrt{n}} = \frac{1}{\sigma\,\sqrt{n}} \sum_{i=1}^{n} (X_i - \mu)$$

وبسبب تطابق التوزيع للمتغيِّرات العشوائية $X_n$ مع توزيع $X$ سينتج أنَّ لجميع المتغيِّرات العشـوائية $X_n^* = X_n - \mu$ توزيع متطابق مع توزيع المتغيِّر العشوائي $X^* = X - \mu$ ، وكذلك لدينا:

$$\mathbf{E}X_n^* = \mathbf{E}X^* = 0 \qquad \& \qquad \mathbf{var}\,X_n^* = \mathbf{var}\,X^* = \sigma^2 \qquad ;\forall n \in \mathbb{N}$$

ومن ثَمّ دوال توزيع المتغيِّرات العشوائية $X_n^*$ ستتطابق مع دالَّة التوزيع $F_{X^*}$ من أجل كل $n \in \mathbb{N}$ ، والتي لها دالَّة مميِّزة $\varphi_{X^*}$ نفسها كالتي للمتغيِّرات العشوائية $X_n^*$ من أجل كل $n \in \mathbb{N}$ أيضاً، ومن ثَمّ يمكننا أن نكتب الآتي:

$$\varphi_{Z_n}(t) = \left(\varphi_{X^*}\left(\frac{t}{\sigma\,\sqrt{n}}\right)\right)^n \qquad ;\forall t \in \mathbb{R}$$

وبنشر الدالَّة $\varphi_{X^*}(t)$ في جوار الصفر وفقاً لـ لنشر ماكلوران Maclaurin Expansion نجد:

$$\varphi_{X^*}(t) = 1 - \frac{1}{2}\,\sigma^2\, t^2 + \theta\left(t^2\right) \qquad ;\forall t \in \mathbb{R}$$

وبالتعويض في صيغة $\varphi_{Z_n}(t)$ يصبح لدينا ما يلي:

$$\varphi_{Z_n}(t) = \left[1 - \frac{t^2}{2n} + \theta\left(\frac{t^2}{n}\right)\right]^n \qquad ;\forall t \in \mathbb{R}$$

علماً أنَّه من أجل كل $t \in \mathbb{R}$ لدينا $\lim\limits_{n\to\infty} n\cdot\theta\left(\frac{1}{n}t^2\right)=0$ ، ومن ثمَّ بوضع: $\tilde{\theta} =: -\frac{1}{2n}t^2+\theta\left(\frac{1}{n}t^2\right)$ نجد أنَّ:

$$\log \varphi_{Z_n}(t) = n\cdot\log\left(1+\tilde{\theta}\right) = n\cdot\left[-\frac{t^2}{2n}+\theta\left(\frac{t^2}{n}\right)\right] = -\frac{t^2}{2}+n\cdot\theta\left(\frac{t^2}{n}\right)$$

وبأخذ نهاية الطرفين عندما $n$ تسعى إلى اللانهاية يصبح لدينا:

$$\lim_{n\to\infty} \log \varphi_{Z_n}(t) = -\frac{t^2}{2}$$

والتي ينتج عنها أنَّ:

$$\lim_{n\to\infty} \varphi_{Z_n}(t) = \mathbf{e}^{-\frac{t^2}{2}}$$

ولكن الطرف الأيمن من العلاقة الأخيرة هي الدالَّة المميِّزة لمتغيِّر عشوائي خاضع للتوزيع الطبيعي المعيّاري، ومن ثمَّ باستخدام مبرهنة الاستمرار لـ ليفي (٣, ٣, ٤, ٧) ينتج لدينا أنَّ:

$$\lim_{n\to\infty} F_{Z_n}(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} \mathbf{e}^{-\frac{t^2}{2}}\, dt \qquad ;\forall z \in \mathbb{R}$$

وبهذا يتمّ البرهان.

**(٣, ٢, ٣, ٨) تطبيق**

لنأخذ $z_1$ و $z_2 \in \mathbb{R}$ مع $z_1 < z_2$، ولنفترض أنَّه من أجل كل $n \in \mathbb{N}$ لدينا $Z_n$ مُعطاة كما في العلاقة [8,20]، فعندئذ يكون لدينا:

$$\lim_{n\to\infty} P\left(z_1 < Z_n < z_2\right) = \lim_{n\to\infty}\left(F_{Z_n}(z_2)-F_{Z_n}(z_1)\right) = \frac{1}{\sqrt{2\pi}}\int_{z_1}^{z_2} \mathbf{e}^{-\frac{t^2}{2}}\, dt$$

ولكن لدينا:

$$P\left(z_1 < Z_n < z_2\right) = P\left(z_1 < \frac{S_n - n\cdot\mu}{\sigma\sqrt{n}} < z_2\right) = P\left(z_1\sigma\sqrt{n}+n\cdot\mu < S_n < z_2\sigma\sqrt{n}+n\cdot\mu\right)$$

وبوضع:

$$s_1 = z_1\sigma\sqrt{n}+n\cdot\mu \qquad \& \qquad s_2 = z_2\sigma\sqrt{n}+n\cdot\mu \qquad\qquad (**)$$

يصبح لدينا ما يلي:

$$\lim_{n\to\infty} P\left(s_1 < S_n < s_2\right) = \frac{1}{\sqrt{2\pi}}\int_{z_1}^{z_2} \mathbf{e}^{-\frac{t^2}{2}}\, dt$$

وهذا يعني أنَّه من أجل $n$ كبيرة بقدر كاف يمكننا أن نكتب:

$$P\left(s_1 < S_n < s_2\right) \approx \frac{1}{\sqrt{2\pi}}\int_{z_1}^{z_2} \mathbf{e}^{-\frac{t^2}{2}}\, dt$$

علماً أنَّ $z_1$ و $z_2$ تُحسبان بوساطة العلاقتين في (**).

**(٨,٣,٢,٤) ملاحظات:**

١- لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية فوق فضاء احتمالي $[\,\Omega,\mathcal{A},P\,]$ مستقلّة ومتطابقة بالتوزيع، وتملك عزوماً ابتدائية من المرتبة الثانية. عندئذ يُقال عن المتغيِّرات العشوائية $S_n$ (المُعطاة في هذه الفقرة) إنَّها **مقاربة للتوزيع الطبيعي** بمتوسط $n\cdot\mu$ وانحراف معياري $\sigma\sqrt{n}$ علماً أنَّ $n\in\mathbb{N}$، وسنكتب ذلك اصطلاحاً بالشكل $S_n \overset{\infty}{\sim} \mathbf{N}\left(n\mu\,,\,\sigma\sqrt{n}\right)$.

٢- إذا كان المجموع لمتغيِّرات عشوائية مستقلّة **مقارب للتوزيع الطبيعي**، فعندئذ يُقال إنَّ نظرية النهاية المركزية مُحقَّقة من أجل هذا المجموع، ومن ثمَّ فإنَّ المتغيِّرات العشوائية $S_n$ لكل $n\in\mathbb{N}$ (والمُعطاة سابقاً في هذه الفقرة) ستتحقَّق من أجلها نظرية النهاية المركزية.

**(٨,٣,٢,٥) أمثلة:**

١- لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية مستقلّة ومتطابقة في التوزيع مع توزيع متغيِّر عشوائي برنولي $X$ بمَعْلَمة $0<p<1$، أي إنَّه من أجل كل $n\in\mathbb{N}$ لدينا ما يلي:

$$P\left(X_n=1\right)=P\left(X=1\right)=p \qquad \& \qquad P\left(X_n=0\right)=P\left(X=0\right)=1-p$$

فعندئذ سيكون المتغيِّر العشوائي $S_n=\sum_{i=1}^{n}X_i$ خاضع للتوزيع الحدَّاني بمَعْلمتين $n$ و$p$ من أجل كل $n\in\mathbb{N}$، ومن جهة أخرى لدينا من أجل كل $n\in\mathbb{N}$ الآتي مُحقَّقاً:

$$\mathbf{E}X_n=p \qquad \& \qquad \mathbf{var}\,X_n=p\,(1-p)$$

أي إنَّه من أجل كل $n\in\mathbb{N}$ سيكون لـ $X_n$ تباين وتوقع رياضياتي منتهٍ ومن ثمَّ تكون شروط استخدام نظرية النهاية المركزية لـ ليندنبرغ- ليفي من أجل المتغيِّرات العشوائية $S_n$ مُحقَّقاً، علماً أنَّه لدينا $\left(S_n \overset{\infty}{\sim} \mathbf{N}\left(np\,,\sqrt{np\,(1-p)}\right)\right)$، ومن هذه النتيجة يتبيَّن لنا أنَّ مسألة موافير- لابلاس ما هي إلاَّ حالة خاصة من مسألة ليندبرغ - ليفي.

٢- لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية مستقلّة ومتطابقة التوزيع مع توزيع متغيِّر عشوائي بواسوني $X$ مَعْلَمته $0<\lambda$ فعندئذ من أجل كل $n\in\mathbb{N}$ سيكون لدينا $\mathbf{E}X_n=\lambda$ و $\mathbf{var}\,X_n=\lambda$، وكذلك:

$$P\,(X_n=k\,)=\frac{\lambda^k}{k\,!}\mathbf{e}^{-\lambda} \quad ;\, k\in\mathbb{N}^o$$

ولنأخذ $S_n$ كما في العلاقة [8,8]، فيكون لدينا:

$$\mathbf{E}S_n=\sum_{i=1}^{n}\mathbf{E}X_i=n\cdot\lambda \qquad \& \qquad \mathbf{var}\,S_n=\sum_{i=1}^{n}\mathbf{var}\,X_i=n\cdot\lambda$$

ولنقم بحساب الاحتمال $P\left(190<S_{100}<210\right)$ من أجل $\lambda=2$.

**الحـل**: بملاحظة أنَّ فرضيات مبرهنة ليندنبرغ-ليفي مُحقَّقة من أجل المسألة المُعطاة، ولذلك يمكننا أن نكتب الآتي:

$$P\left(190<S_{100}<210\right)=P\left(\frac{190-200}{10\sqrt{2}}<\frac{S_{100}-200}{10\sqrt{2}}<\frac{210-200}{10\sqrt{2}}\right)=P\left(-0.71<Z_{100}<0.71\right)$$
$$\approx\Phi(0.71)-\Phi(-0.71)=0.7611-0.2389=0.5222$$

علماً أنَّ $Z_{100}:=\dfrac{S_{100}-200}{10\sqrt{2}}$، وبهذا يتمُّ الحل.

تعرض لنا النتيجة الآتية إحدى أهم النتائج لنظرية ليندنبرغ - ليفي.

**(٨,٣,٢,٦) نتيجة**

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$ مستقلّة ومتطابقة بالتوزيع مع توزيع متغيِّر عشوائي $X$، وعناصرها تملك عزوماً ابتدائيـة من المرتبة الثانية مع:

$$\mathbf{E}X_n = \mathbf{E}X = \mu \qquad \& \qquad \mathbf{var}X_n = \mathbf{var}X = \sigma^2 > 0$$

ولنأخذ متتالية المتغيِّرات العشوائية $(\bar{X}_n)_{n\in\mathbb{N}}$ علماً أنَّ $\bar{X}_n$ مُعطى بالعلاقة [8,11-a]، فعندئذ تكون المتتالية $(\bar{X}_n)_{n\in\mathbb{N}}$ **مقاربة للتوزيع الطبيعي** بمتوسط $\mu$ وانحراف معياري $\frac{\sigma}{\sqrt{n}}$، أي إنَّ $\bar{X}_n \overset{\infty}{\sim} \mathbf{N}\left(\mu\,,\frac{\sigma}{\sqrt{n}}\right)$.

**الإثبات**: بملاحظة أنَّه من أجل كل $n \in \mathbb{N}$ يمكننا أن نكتب الآتي:

$$\mathbf{E}\bar{X}_n = \frac{n\,\mu}{n} = \mu \qquad \& \qquad \mathbf{var}\,\bar{X}_n = \frac{n\cdot\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

وكذلك بوضع:

$$Z_n = \frac{\bar{X}_n - \mathbf{E}\bar{X}_n}{\sqrt{\mathbf{var}\,\bar{X}_n}} = \frac{\frac{1}{n}\sum_{i=1}^{n} X_i - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\sum_{i=1}^{n} X_i - n\,\mu}{\sigma\sqrt{n}}$$

فإنَّه بحسب الملاحظة / ١/ من (٨,٣,٢,٤) ستكون المتتالية $(\bar{X}_n)_{n\in\mathbb{N}}$ **مقاربة للتوزيع الطبيعي** بمتوسط $\mu$ وانحراف معياري $\frac{\sigma}{\sqrt{n}}$، وبهذا يتمّ الإثبات.

**(٨,٣,٢,٧) أمثلة**

١ – لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية مستقلّة ومتطابقة بالتوزيع مع متغيِّر عشوائي $X$ خاضع للتوزيع المنتظم المتقطِّع بقيم في $\mathbb{N}_9^o$، ولننظر ما الذي يحدث لقيمة الاحتمال المتعلّق بالمتوسط $\bar{X}_n$ عندما نبتعد عن القيمة المتوقعة لـ $X_n$ وذلك من أجل قيم مثبَّتة لـ $n \in \mathbb{N}$.

**الحـل**: لدينا من أجل كل $n \in \mathbb{N}$ ما يلي:

$$P(X = k) = 0.1 \qquad ;\forall\ k \in \mathbb{N}_9^o$$

وكذلك:

$$\mathbf{E}X_n = \frac{1}{10}\sum_{k=0}^{9} k = 4.5 = \mu \qquad \& \qquad \mathbf{var}X_n = \frac{1}{10}\sum_{k=0}^{9}(k-\mu)^2 = 8.25 = \sigma^2$$

وبناءً على النتيجة السابقة يكون لدينا $\bar{X}_n \overset{\infty}{\sim} \mathbf{N}\left(4.5\,,\frac{2.87}{\sqrt{n}}\right)$، فلو أخذنا على سبيل المثال قيمة مُثبَّة لـ $n$ ولتكن $n = 25$، فعندئذ سيكون $\bar{X}_{25} \overset{\infty}{\sim} \mathbf{N}\left(4.5\,,\frac{2.87}{5}\right)$، ومن ثمَّ يمكننا أن نكتب:

$$\begin{aligned} P\left(\bar{X}_{25} > 5\right) &= P\left(\frac{\bar{X}_{25}-4.5}{0.574} > \frac{5-4.5}{0.574}\right) = 1 - P\left(\frac{\bar{X}_{25}-4.5}{0.574} \le 0.87\right) \\ &\approx 1 - \Phi(0.87) = 1 - 0.8078 = 0.1922 \end{aligned}$$

وبالمثل نجد أنَّ:

$$P\left(\bar{X}_{25} > 5.5\right) = P\left(\frac{\bar{X}_{25} - 4.5}{0.574} > \frac{5.5 - 4.5}{0.574}\right) = 1 - P\left(\frac{\bar{X}_{25} - 4.5}{0.574} \le 1.74\right)$$
$$\approx 1 - \Phi(1.74) = 1 - 0.9591 = 0.04090$$

وهكذا فأنَّه من أجل قيمة مُثبَّة $n \in \mathbb{N}$ نلاحظ أنَّه كلما ابتعدت قيمة المتوسط $\bar{X}_n$ عن القيمة المتوقعة لـ $X$، تناقصت قيمة الاحتمال للحادث المدروس، وإحصائياً لو كانت $\bar{x}_n$ (قيمة $\bar{X}_n$) هي متوسط قيم بيانات عينة حجمها $n$، وكان $X$ واصفاً لمجتمع إحصائي متوسطه $\mu$، فإنَّ أفضل النتائج لهذه العينة نحصل عليها من أجل $\mu = \mathbf{E}X = \bar{x}_n$ .

الآن لنأخذ قيمة أخرى لـ $n$ (مع ثبات قيمة الانحراف المعياري والمتوسط) ولننظر إن كان ذلك سيؤدي إلى تغيّر في النتائج أم لا؟

أ- من أجل $n = 64$ لدينا $\bar{X}_{64} \overset{\infty}{\sim} \mathbf{N}\left(4.5\ , \frac{2.87}{8}\right)$، ومن ثمَّ يمكننا أن نكتب:

$$P\left(\bar{X}_{64} > 5\right) = P\left(\frac{\bar{X}_{64} - 4.5}{0.359} > \frac{5 - 4.5}{0.359}\right) = 1 - P\left(\frac{\bar{X}_{64} - 4.5}{0.359} \le 1.39\right)$$
$$\approx 1 - \Phi(1.39) = 1 - 0.9177 = 0.0823$$

وبالمثل نجد أنَّ:

$$P\left(\bar{X}_{64} > 5.5\right) = P\left(\frac{\bar{X}_{64} - 4.5}{0.359} > \frac{5.5 - 4.5}{0.359}\right) = 1 - P\left(\frac{\bar{X}_{64} - 4.5}{0.359} \le 2.89\right)$$
$$\approx 1 - \Phi(2.89) = 1 - 0.9981 = 0.0019$$

فنلاحظ أنَّ قيم الاحتمالات قد تناقصت بكبر قيمة $n$.

ب- لنأخذ الآن قيمة $n$ تساوي 100، فعندئذ يكون $\bar{X}_{100} \overset{\infty}{\sim} \mathbf{N}\left(4.5\ , \frac{2.87}{10}\right)$، ومن ثمَّ يمكننا أن نكتب:

$$P\left(\bar{X}_{100} > 5\right) = P\left(\frac{\bar{X}_{100} - 4.5}{0.287} > \frac{5 - 4.5}{0.287}\right) = 1 - P\left(\frac{\bar{X}_{100} - 4.5}{0.287} \le 1.74\right)$$
$$\approx 1 - \Phi(1.74) = 1 - 0.9591 = 0.04090$$

وبالمثل نجد أنَّ:

$$P\left(\bar{X}_{100} > 5.5\right) = P\left(\frac{\bar{X}_{100} - 4.5}{0.287} > \frac{5.5 - 4.5}{0.287}\right) = 1 - P\left(\frac{\bar{X}_{100} - 4.5}{0.287} \le 3.48\right)$$
$$\approx 1 - \Phi(3.48) = 1 - 0.9997 = 0.00030$$

وهكذا نلاحظ أنَّ احتمال ابتعاد قيمة المتوسط عن القيمة المتوقعة لـ $X_n$ تناقصت أكثر فأكثر كلما كبرت قيمة $n \in \mathbb{N}$ أيضاً، ومن ثمّ فكلما كَبُرت قيمة $n \in \mathbb{N}$ فإنَّ ابتعاد قيمة المتوسط $\bar{X}_n$ عن القيمة المتوقعة لـ $X$ تؤدي إلى تناقص قيمة الاحتمال للحادث المدروس أكثر فأكثر، وإحصائياً يعني ذلك أنَّه كلما كَبُر حجم العينة $n \in \mathbb{N}$، تكشفت لنا حقائق ومعلومات إضافية تقربنا أكثر من سلوك المجتمع المدروس، ومن ثمَّ تكون نتائج دراسة هذه العينة أفضل.

٢- لتكن $(X_n)_{n \in \mathbb{N}}$ متتالية متغيِّرات عشوائية مستقلّة ومتطابقة التوزيع مع متغيِّر عشوائي $X$ خاضع للتوزيع المنتظم المستمر على الفترة $[0,1]$، والمطلوب حساب احتمال أن يكون للمتوسط $\bar{X}_n$ قيمة أصغر من $\mathbf{E}X = \mu$، وذلك من أجل قيم مختلفة لـ $n \in \mathbb{N}$.

الحـل: لدينا من أجل كل $n \in \mathbb{N}$ القيم الآتية:

$$\mathbf{E}X_n = \mathbf{E}X = \frac{1}{2} = \mu \qquad \& \qquad \mathbf{var}X_n = \mathbf{var}X = \frac{1}{12}$$

وبناءً على النتيجة السابقة يكون $\bar{X}_n \overset{\infty}{\sim} \mathbf{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$، فلو أخذنا $n = 49$ فعندئذ يكون لدينا:

$$\bar{X}_{49} = \frac{X_1 + X_2 + \dots + X_{49}}{49}$$

ومن ثمَّ يكون لدينا $\bar{X}_{49} \overset{\infty}{\sim} \mathbf{N}\left(\frac{1}{2}, \frac{0.2887}{7}\right)$، ولنقم بحساب الاحتمالين الآتيين:

$$P\left(\bar{X}_{49} < 0.45\right) = ? \qquad \& \qquad P\left(\bar{X}_{49} < 0.40\right) = ?$$

لحساب الاحتمال $P\left(\bar{X}_{49} < 0.45\right)$ نجد ما يلي:

$$\begin{aligned} P\left(\bar{X}_{49} < 0.45\right) &= P\left(\frac{\bar{X}_{49} - 0.5}{0.041} < \frac{0.45 - 0.5}{0.041}\right) \\ &= P\left(\frac{\bar{X}_{49} - 0.5}{0.041} < -1.22\right) \approx \Phi(-1.22) \approx 0.1112 \end{aligned}$$

وأمّا لحساب الاحتمال $P\left(\bar{X}_{49} < 0.40\right)$ نجد بأسلوب مماثل لما سبق ما يلي:

$$P\left(\bar{X}_{49} < 0.4\right) = P\left(\frac{\bar{X}_{49} - 0.5}{0.041} < \frac{0.4 - 0.5}{0.041}\right) \approx \Phi(-2.43) \approx 0.0075$$

وأمّا لو أخذنا الآن $n = 96$ فعندئذ سيكون لدينا $\bar{X}_{96} \overset{\infty}{\sim} \mathbf{N}\left(\frac{1}{2}, \frac{0.2887}{\sqrt{96}}\right)$، ومن ثمَّ يكون لدينا:

$$\begin{aligned} P\left(\bar{X}_{96} < 0.45\right) &= P\left(\frac{\bar{X}_{96} - 0.5}{0.0295} < \frac{0.45 - 0.5}{0.0295}\right) \\ &= P\left(\frac{\bar{X}_{96} - 0.5}{0.0295} < -1.69\right) \approx \Phi(-1.69) \approx 0.0455 \end{aligned}$$

وأمّا لحساب الاحتمال $P\left(\bar{X}_{96} < 0.40\right)$ فنجد بأسلوب مماثل لما سبق ما يلي:

$$P\left(\bar{X}_{96} < 0.4\right) = P\left(\frac{\bar{X}_{96} - 0.5}{0.0295} < \frac{0.4 - 0.5}{0.0295}\right) \approx \Phi(-3.39) = 0.0003$$

وهكذا نلاحظ أنَّ المناقشة التي قدِّمت في المثال السابق (**بخصوص تناقص قيمة الاحتمال كلما ابتعدت قيمة المتوسط عن القيمة المتوقعة لـ $X_n$ مع** $n \in \mathbb{N}$) تبقى سارية المفعول من أجل هذا المثال أيضاً.

### (٩,٢,٣,٨) ملاحظات

١- في الواقع إنَّ النتيجة السابقة (**بخصوص تناقص قيمة الاحتمال كلما ابتعدت قيمة المتوسط عن القيمة المتوقعة لـ $X_n$**) تبقى سارية المفعول على كل المتغيِّرات العشوائية التي تملك عزوماً ابتدائية من المرتبة الثانية **عدا** المتغيِّرات العشوائية الخاضعة للتوزيع الوحيد النقطة.

٢- إذا كانت $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية مستقلّة ومتطابقة في التوزيع ولكنَّها **لا تملك** عزوماً ابتدائية فعندئذ تصبح مبرهنة **ليندنبرغ–ليفي** غير قابلة للتطبيق على هذه المتتالية، ومن ثمَّ يجب البحث في منحى آخر للتعامل مع توزيع المتوسط $\bar{X}_n$ (**غالباً باستخدام الدوال المميِّزة**).

### (٨,٣,٣) مسألة ليابانوف Lyapunov Problem

لقد بينّا سابقاً من خلال مسألة موافير– لابلاس و ليندنبرغ – ليفي أنَّه إذا كانت $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية مستقلّة ومتطابقة في التوزيع، فإنَّ دالَّة توزيع المجموع لهذه المتغيِّرات العشوائية سوف تسعى إلى دالَّة التوزيع الطبيعي المعياري. لكن إذا كانت هذه المتغيِّرات العشوائية ليست متطابقة في التوزيع، فعندئذ ليس بالضرورة أن تسعى دالَّة توزيع المجموع لهذه المتغيِّرات العشوائية إلى التوزيع الطبيعي، ولكن إذا فُرضت شروط مُحَدَّدة على متتالية المتغيِّرات العشوائية المستقلّة فحينئذ قد يصبح الوصل إلى تقريب للتوزيع الطبيعي ممكناً. إنَّ نظرية ليابانوف الآتية تقدِّم لنا الشروط الكافية لكي يتقارب مجموع متغيِّرات عشوائية مستقلّة (حتى لو كانت غير متطابقة في التوزيع) من التوزيع الطبيعي.

### (٨,٣,٣,٢) نظرية ليابانوف Lyapunov Theorem

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية مستقلّة، وتملك عزوماً ابتدائية من المرتبة الثالثة، وبفرض أنَّ:

$$B_n := \sqrt[3]{\sum_{k=1}^{n} \beta_k} \qquad \& \qquad C_n := \sqrt{\sum_{k=1}^{n} \sigma_k^2} \qquad \& \qquad Z_n := \frac{\sum_{k=1}^{n} (X_k - \mu_k)}{C_n}$$

علماً أنَّه من أجل كل $k \in \mathbb{N}$ لدينا $\mu_k = \mathbf{E}X_k$ و $\sigma_k^2 = \mathbf{var}X_k > 0$.

الآن إذا كانت $\lim\limits_{n\to\infty} \frac{B_n}{C_n}$ موجودةً وتحقِّق العلاقة الآتية:

$$\lim_{n\to\infty} \frac{B_n}{C_n} = 0 \qquad [8,22]$$

فعندئذ من أجل كل $z \in \mathbb{R}$ ستكون العلاقة الآتية مُحَقَّقة:

$$\lim_{n\to\infty} F_{Z_n}(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathbf{e}^{-\frac{t^2}{2}}\, dt$$

**البرهان:** على سبيل تبسيط العمل سنضع من أجل كل $k \in \mathbb{N}$ ما يلي:

$$X_k^* := X_k - \mu_k \qquad \& \qquad Y_k := \frac{X_k^*}{C_k}$$

وبما أنَّ $\mathbf{E}X_k^* = 0$ وكذلك العزوم $\sigma_k^2$ و $\mu_k^{(3)}$ موجودة، فإنَّه يمكن كتابة الآتي:

$$\varphi_{X_k^*}(t) = 1 - \frac{\sigma_k^2\, t^2}{2} + \frac{\mu_k^{(3)}}{6}(\mathbf{i}\,t)^3 + \theta\left(\mu_k^{(3)} t^3\right)$$

ومن ثَمّ فإنَّه من أجل كل $t \in \mathbb{R}$ سيكون:

$$\varphi_{Y_k}(t) = \varphi_{X_k^*}\left(\frac{t}{C_n}\right) = 1 \underbrace{- \frac{\sigma_k^2\, t^2}{2C_n^2} + \frac{\mu_k^{(3)}}{6C_n^3}(\mathbf{i}\,t)^3 + \theta\left(\frac{\mu_k^{(3)}\, t^3}{C_n^3}\right)}_{u_{k,n}} = 1 + u_{k,n}$$

علماً أنَّه من أجل كل عدد منته $t \in \mathbb{R}$ لدينا ما يلي:

$$\lim_{n\to\infty} \frac{C_n^3}{\mu_k^{(3)} t^3}\, \theta\left(\frac{\mu_k^{(3)} t^3}{C_n^3}\right) = 0 \qquad (1)$$

ولكن بحسب متباينة ليابانوف العلاقة [20–7] من / ٢٣ , ١ , ٢ , ٧/ سيكون $\sigma_k \le \sqrt[3]{\beta_k^{(3)}}$ ، ومن ثَمّ بحسب الشرط [22–8] ينتج أنَّه من أجل قيمة منتهية $t \in \mathbb{R}$ ستكون المتباينة الآتية مُحقَّقة:

$$\lim_{n\to\infty}\left|\frac{-\sigma_k^2\cdot t^2}{2C_n^2}\right| \le \lim_{n\to\infty}\frac{\sqrt[3]{\left(\beta_k^{(3)}\right)^2}}{2C_n^2}t^2 \le \lim_{n\to\infty}\frac{1}{2}\left(\frac{B_n^2}{C_n^2}t^2\right)=0$$

ومن ثَمَّ يصبح لدينا:

$$\lim_{n\to\infty}\left|\frac{\mu_k^{(3)}}{6C_n^3}(\mathrm{i}\,t)^3\right| \le \lim_{n\to\infty}\frac{\beta_k^{(3)}}{6C_n^3}|t|^3 \le \lim_{n\to\infty}\frac{B_n^3}{6C_n^3}|t|^3=0$$

ومنه ينتج أنّ $\lim_{n\to\infty} u_{k,n}=0$ ، وبسبب التقارب المنتظم لمتتالية الدوال $\left(u_{k,n}\right)_{n\in\mathbb{N}}$ بالنسبة إلى $k$، فإنَّه من أجل كل $t \in \mathbb{R}$ مثَّبتة يوجد عدد طبيعي $N=N(t)$ بحيث إنَّه من أجل كل القيم $n>N$ وكل $n\ge k$ يكون لدينا $\left|u_{k,n}\right|<\frac{1}{2}$ ، ومن ثَمَّ نحصل على العلاقة الآتية:

$$\begin{aligned}\log\varphi_{Y_k}(t) = \log\left(1+u_{k,n}\right) &= \frac{u_{k,n}}{1}-\frac{u_{k,n}^2}{2}+\frac{u_{k,n}^3}{3}-\frac{u_{k,n}^4}{4}+\ldots+(-1)^{\ell}\frac{u_{k,n}^{\ell}}{\ell}+\ldots\\ &= u_{k,n}-\frac{u_{k,n}^2}{2}\left(1-\frac{2}{3}u_{k,n}+\frac{2}{4}u_{k,n}^2-\ldots\right)\end{aligned}$$

وبوضع $V_{k,n}:=1-\frac{2}{3}u_{k,n}+\frac{2}{4}u_{k,n}^2-\ldots$ تنتج المساواة الآتية:

$$\log\varphi_{Y_k}(t)=u_{k,n}-\frac{u_{k,n}^2}{2}V_{k,n}$$

ولكن بسبب أنَّ $\left|u_{k,n}\right|<\frac{1}{2}$ فإنَّه يمكن كتابة الآتي:

$$\left|V_{k,n}\right| \le 1+\frac{2}{3}\left|u_{k,n}\right|+\frac{2}{4}\left|u_{k,n}\right|^2+\ldots<1+\left|u_{k,n}\right|+\left|u_{k,n}\right|^2+\ldots<1+\frac{1}{2}+\left(\frac{1}{2}\right)^2+\left(\frac{1}{2}\right)^3+\ldots=\sum_{s=0}^{\infty}\left(\frac{1}{2}\right)^s=2$$

وبوضع $U_{k,n}:=-\frac{V_{k,n}}{2}$ ينتج أنَّ $\left|U_{k,n}\right|<1$ ، ومن ثَمّ يكون لدينا:

$$\log\varphi_{Y_k}(t)=u_{k,n}+u_{k,n}^2\,U_{k,n}$$

والآن من أجل كل $n\in\mathbb{N}$ سنضع $Z_n=\sum_{k=1}^{n}Y_k$ ، وبسبب الاستقلال بين المتغيِّرات العشوائية $Y_1$ ،$Y_2$ ،…و $Y_n$ فإنَّه من أجل كل $t\in\mathbb{R}$ ستكون العلاقة الآتية مُحقَّقة:

$$\varphi_{Z_n}(t)=\prod_{k=1}^{n}\varphi_{Y_k}(t) \qquad (2)$$

ومنه ينتج لدينا:

$$\log\varphi_{Z_n}(t)=\sum_{k=1}^{n}\log\varphi_{Y_k}(t)=\sum_{k=1}^{n}\left(u_{k,n}+u_{k,n}^2\,U_{k,n}\right)$$

وإضافة لذلك لدينا:

$$\sum_{k=1}^{n} u_{k,n} = -\frac{t^2}{2} + \sum_{k=1}^{n} \frac{\mu_k^{(3)}(\mathbf{i}\,t)^3}{6C_n^3} + \sum_{k=1}^{n} \theta\left(\frac{\mu_k^{(3)} t^3}{C_n^3}\right) \qquad \textbf{(3)}$$

ولكن من أجل أي عدد منتهٍ $t \in \mathbb{R}$ لدينا:

$$\lim_{n\to\infty}\left|\sum_{k=1}^{n}\frac{\mu_k^{(3)}(\mathbf{i}\,t)^3}{6C_n^3}\right| \le \lim_{n\to\infty}\sum_{k=1}^{n}\frac{\beta_k^{(3)}|t|^3}{6C_n^3} = \lim_{n\to\infty}\frac{B_n^3|t|^3}{6C_n^3} = 0 \qquad \textbf{(4)}$$

وباستخدام العلاقة **(1)** نحصل على العلاقة الآتية:

$$\lim_{n\to\infty}\sum_{k=1}^{n}\theta\left(\frac{\mu_k^{(3)}t^3}{6C_n^3}\right) \le \lim_{n\to\infty}\sum_{k=1}^{n}\left[\frac{\mu_k^{(3)}t^3}{6C_n^3}\cdot\frac{6C_n^3}{\alpha_k t^3}\cdot\theta\left(\frac{\mu_k^{(3)}t^3}{6C_n^3}\right)\right] = 0$$

ومن هذه العلاقة الأخيرة ومراعاة العلاقتين العلاقة **(3)** و**(4)** ينتج من أجل أي عدد منتهٍ $t \in \mathbb{R}$ أنّ العلاقة الآتية صحيحة:

$$\lim_{n\to\infty}\sum_{k=1}^{n} u_{k,n} = -\frac{t^2}{2} \qquad \textbf{(5)}$$

والآن من أجل حساب نهاية المجموع الآتي:

$$\sum_{k=1}^{n} u_{k,n}^2 = \sum_{k=1}^{n}\left[-\frac{\sigma_k^3 t^2}{2C_n^2} + \frac{\mu_k^{(3)}(\mathbf{i}\,t)^3}{6C_n^3} + \theta\left(\frac{\mu_k^{(3)}t^3}{C_n^3}\right)\right]^2$$

لدينا من متباينة ليابانوف ومراعاة العلاقة [22–8] ما يلي محقّقاً:

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{\sigma_k^4 t^4}{4C_n^4} \le \lim_{n\to\infty}\sum_{k=1}^{n}\frac{\sqrt[3]{\left(\beta_k^{(3)}\right)^4}}{4C_n^4}\cdot t^4 = \lim_{n\to\infty}\sum_{k=1}^{n}\frac{\beta_k^{(3)}\sqrt[3]{\beta_k^{(3)}}}{4C_n^3 C_n}\cdot t^4$$

$$\le \lim_{n\to\infty}\sum_{k=1}^{n}\frac{\beta_k^{(3)}}{4C_n^3}\cdot t^4 = \lim_{n\to\infty}\sum_{k=1}^{n}\frac{B_n^3}{4C_n^3}\cdot t^4 = 0$$

وكذلك:

$$\lim_{n\to\infty}\left|\sum_{k=1}^{n}\left[\frac{\mu_k^{(3)}t^3}{6C_n^3}\right]^2\right| \le \lim_{n\to\infty}\sum_{k=1}^{n}\left|\frac{\mu_k^{(3)}t^3}{6C_n^3}\right|^2 \le \lim_{n\to\infty}\sum_{k=1}^{n}\frac{\left(\beta_k^{(3)}\right)^2 t^6}{36\,C_n^6} \le \lim_{n\to\infty}\sum_{k=1}^{n}\frac{B_n^6 t^6}{36\,C_n^6} = 0$$

وبعد الأخذ بالحسبان العلاقة **(1)** نجد:

$$\lim_{n\to\infty}\sum_{k=1}^{n}\left(\theta\left(\frac{\mu_k^{(3)}t^3}{6C_n^3}\right)\right)^3 = 0$$

وكذلك سنحصل على العلاقات الثلاث الآتية:

$$\lim_{n\to\infty}\left|\sum_{k=1}^{n-1}\sum_{j=k+1}^{n}\frac{t^2\mu_k^{(3)}}{C_n^2}\,\theta\left(\frac{\mu_j^{(3)}t^3}{C_n^3}\right)\right| = 0 \qquad \& \qquad \lim_{n\to\infty}\left|\sum_{k=1}^{n-1}\sum_{j=k+1}^{n}\frac{\left(\mu_k^{(3)}\right)^2\mu_j^{(3)}\,\mathbf{i}^3 t^5}{6\,C_n^5}\right| = 0$$

$$\lim_{n\to\infty}\left|\sum_{k=1}^{n-1}\sum_{j=k+1}^{n}\frac{\mu_k^{(3)}\,\mathbf{i}^3 t^5}{3\,C_n^3}\,\theta\left(\frac{\mu_j^{(3)}t^3}{C_n^3}\right)\right| = 0$$

وبما أنَّ $\left|u_{k,n}\right| < \frac{1}{2}$، وبعد الأخذ بالحسبان العلاقات الست الأخيرة فإنَّه سينتج لدينا الآتي:

$$\lim_{n\to\infty} \sum_{k=1}^{n} U_{k,n}\, u_k^2 = 0 \qquad ;\forall\ k \in \mathbb{N}$$

ومن هذه العلاقة الأخيرة وبعد الأخذ بالحسبان العلاقة **(2)** ومراعاة العلاقة **(5)** فإنَّه سينتج لدينا أنَّه من أجل كلّ $t \in \mathbb{R}$ ما يلي مُحقَّقاً:

$$\lim_{n\to\infty} \log \varphi_{Z_n}(t) = -\frac{t^2}{2}$$

ومن ثَمّ لأجل كلّ $t \in \mathbb{R}$ يصبح لدينا :

$$\lim_{n\to\infty} \varphi_{Z_n}(t) = \mathrm{e}^{-\frac{t^2}{2}}$$

ومن هذه العلاقة الأخيرة ومن مبرهنة ليفي ينتج أنَّه من أجل كل $z \in \mathbb{R}$ ما يلي مُحقَّقاً:

$$\lim_{n\to\infty} F_{Z_n}(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-\frac{t^2}{2}}\, dt$$

وبهذا يتمُّ البرهان.

لقد لاحظنا أنَّ نظرية ليابانوف تعطي شرطاً كافاً فقـط لتحقيق العلاقة [22–8]، ولذلك سنقدم فيما يلي نظرية (سنقبلها دون برهان) تعطينا شرطاً لازماً وكافياً لتَحقُّق العلاقة [22–8] ، وهذه النظرية تُعرف باسم مسألة ليندبرغ – فيلر .

## (٨,٣,٤) مسألة ليندنبرغ – فيلر Lindenberg – Feller Problem

إنَّ هذه المسألة تنطلق من على مستو جيد من العمومية مقارنة بنظريات النهاية المركزية السابقة، إذ إنَّها تنطلق من أسرة متغيِّرات عشوائية مستقلّة، وليس بالضرورة أن تكّون متطابقةً في التوزيع، وتُبيِّن أنَّه تحت شروط مُحدَّدة يمكن الوصول إلى تقريب للتوزيع الطبيعي، وهذا هو محتوى النظرية الآتية.

### (٨,٣,٤,١) نظرية ليندنبرغ – فيلر Lindenberg – Feller Theorem

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية فوق فضاء احتمالي $[\,\Omega, \mathscr{A}, P\,]$ مستقلّة وتملك عزوماً ابتدائية من المرتبة الثانية، ولنضع من أجل كل $n \in \mathbb{N}$ ما يلي:

$$C_n := \sqrt{\sum_{k=1}^{n} \sigma_k^2} \qquad \& \qquad Z_n := C_n^{-1} \cdot \sum_{k=1}^{n} \left(X_n - \mu_k\right)$$

عندئذ الشرط اللازم والكافي لتحقُّق العلاقتين الآتيتين:

$$\text{a)}\quad \lim_{n\to\infty} \max_{1\le k\le n} \frac{\sigma_k}{C_n} = 0 \qquad \& \qquad \text{b)}\quad \lim_{n\to\infty} F_{Z_n}(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-\frac{t^2}{2}}\, dt$$

هو أن يكون من أجل كل $\varepsilon > 0$ لدينا العلاقة الآتية مُحقَّقة:

$$\lim_{n\to\infty} \frac{1}{C_n^2} \sum_{k=1}^{n} \int_{\left|x-\mu_k\right| > \varepsilon\cdot C_n} \left(x-\mu_k\right)^2 dF_{X_k}(x) = 0 \qquad [8,23\text{-a}]$$

### (٨,٣,٤,٢) ملاحظة

إذا كان للمتغيِّرات العشوائية $X_n$ دوال كثافة احتمالية $f_{X_n}$ من أجل كل $n \in \mathbb{N}$، فعندئذ تكتب العلاقة [8–23-a] على

النحو الآتي أيضاً:

$$\lim_{n\to\infty} \frac{1}{C_n^2} \sum_{k=1}^{n} \int_{|x-\mu_k|>\varepsilon\cdot C_n} (x-\mu_k)^2 f_{X_k}(x)\,(dx) = 0 \qquad [8,23\text{-}b]$$

وأمّا إذا كانت المتغيِّرات العشوائية $X_n$ متقطِّعة من أجل كلّ $n \in \mathbb{N}$، وكتلها الاحتمالية $p_{n,\ell}$ عند النقاط $x_{n,\ell}$ من أجل كل $\ell \in L \subseteq \mathbb{N}$، فعندئذ تُكتب العلاقة [8–23-a] على النحو الآتي:

$$\lim_{n\to\infty} \frac{1}{C_n^2} \sum_{k=1}^{n} \sum_{|x_{k,\ell}-\mu_k|>\varepsilon\cdot C_n} (x_{k,\ell}-\mu_k)^2\, p_{k,\ell} = 0 \qquad [8,23\text{-}c]$$

من نظرية ليندنبرغ – فيلر السابقة تنتج لدينا المقولة الآتية التي سنوردها دون إثبات أيضاً.

**(٨,٣,٤,٣) نتيجة**

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية مستقلّة وعناصرها **تملك** عزوماً ابتدائية من المرتبة الثانية مع $\mathbf{var}\,X_n > 0$ من أجل كل $n \in \mathbb{N}$، ولنضع بالتعريف ما يلي:

$$C_n = \sqrt{\sum_{k=1}^{n} \sigma_k^2} \qquad \& \qquad Z_n = C_n^{-1}\cdot \sum_{k=1}^{n} (X_n - \mu_k)$$

فإذا وجد ثابت $c > 0$ بحيث يكون من أجل كل $n \in \mathbb{N}$ لدينا $P(|X_n| \le c) = 1$، فعندئذ تحقّق العلاقة $\lim_{n\to\infty} C_n^2 = +\infty$ هو شرط لازم وكافٍ لكي تكون العلاقة الآتية مُحقَّقة:

$$\lim_{n\to\infty} F_{Z_n}(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-\frac{t^2}{2}}\, dt$$

حيث نلاحظ هنا أنَّ تحقُّق العلاقة $\lim_{n\to\infty} C_n^2 = +\infty$ يكافئ تحقُّق العلاقة $\sum_{n=1}^{\infty} \sigma_n^2 = +\infty$.

**(٨,٣,٤,٤) أمثلة**

١– لنأخذ $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية برنوليّة مستقلّة فوق فضاء احتمالي $[\Omega, \mathcal{A}, P]$ بحيث إنَّه من أجل كل $n \in \mathbb{N}$ لـ $X_n$ مَعْلَمة $p_n$، فعندئذ سيكون من أجل كل $k \in \mathbb{N}_n$ لدينا القيم الآتية:

$$\mu_k = \mathbf{E}\,X_k = p_k \qquad \& \qquad \sigma_k^2 = \mathbf{var}\,X_k = p_k\cdot(1-p_k) \qquad \& \qquad C_n^2 = \sum_{k=1}^{n} \sigma_k^2 = \sum_{k=1}^{n} p_k\cdot(1-p_k)$$

وهنا نجد أنَّ تباعد المتسلسلة العددية $\sum_{n=1}^{\infty} p_k\cdot(1-p_k)$ هو شرط لازم وكافٍ حتى يكون توزيع المتغيِّر العشوائي $Z_n$ مقاربًا للتوزيع الطبيعي بمتوسط $\sum_{k=1}^{n} p_k$ وتباين $\sum_{k=1}^{n} p_k\cdot(1-p_k)$، أي إنَّه يكون $Z_n \overset{\infty}{\sim} \mathbf{N}\left(\sum_{k=1}^{n} p_k\,,\sqrt{\sum_{k=1}^{n} p_k\cdot(1-p_k)}\right)$ **إذا وفقط إذا** كانت المتسلسلة العددية $\sum_{k=1}^{\infty} p_k\cdot(1-p_k)$ متباعدة نحو $\infty$، وتجدر الإشارة هنا إلى أنَّه يُقال في هذه الحالة عن المتغيِّر العشوائي $Z_n$ إنَّه خاضع للتوزيع ثنائي النقطة المتعدِّد أو الحدَّاني المُعمَّم.

٢– يوجد في مستودع أنواع مختلفة من المصابيح معبأة في صناديق بحيث يحوي كل صندوق عدداً منتهياً من المصابيح. إنَّ حظنا في الحصول على مصباح معيب في شحنة ليس نفسه في كل الشحنات التي تعود إلى أنواع مختلفة. الآن سنفترض أنَّ كل صندوق يحوي

عدداً كبيراً جداً من المصابيح (ومن ثمَّ إزالة بعض المصابيح منها لا يؤثّر في ترتيبها ووضعها ضمن الصندوق)، وأنَّه لدى إنتاج الصندوق ذي الرقم $k$ يكون $p_k$ هو احتمال الحصول على مصباح سليم (احتمال النجاح) من هذا الصندوق، وأنَّه لدينا:

$$p_1 = 0.95 \;\; \& \; p_2 = 0.90 \;\; \& \;\; p_3 = 0.98 \;\; \& \;\; p_4 = 0.92 \; \& \;\; p_5 = 0.96$$

قمنا بسحب عشوائي لعشرين مصباحاً من كل صندوق، ومن ثمَّ اختبار هذه المصابيح فوجد منها 11 مصباحاً معيباً (لا تحقّق المواصفات المطلوبة)، وهذه النتيجة أثارت في نفسنا الشك تجاه افتراضنا للقيم $p_k$ (فقد تكون تلك القيم مبالغ فيها)، ولكي نبيِّن إن كان شكّنا مبررًا أم لا سنقوم بدراسة هذه القضية على النحو الآتي. إنَّ النموذج الرياضياتي لهذه المسألة يبدو لنا كما يلي:

لدينا 100 متغيِّر عشوائي مستقل $X_1$، $X_2$،...و $X_{100}$ برنوليّة، ولكنَّها ليست جميعها متطابقة في التوزيع. إنَّ هذه المتغيِّرات العشوائية موزعة في خمس فئات متساوية الحجم بحيث إن الفئة ذات الرقم $k$ تنتمي إليها تلك المتغيِّرات العشوائية ذات الاحتمالات $p_k$ من أجل أي $k \in \mathbb{N}_5$، ولنشكّل المجموع الآتي:

$$S_{100} = X_1 + \ldots + X_{20} + X_{21} + \ldots + X_{40} + X_{41} + \ldots + X_{100}$$

فعندئذ يكون لـ $S_{100}$ توزيع حدّاني مُعمَّم بالقيم المميَّزة الآتية:

$$\mathbf{E}S_{100} = (20)(0.95) + (20)(0.90) + (20)(0.98) + (20)(0.92) + (20)(0.96) = 94.20$$

$$\begin{aligned}\mathbf{var}\,S_{100} &= (20)(0.05)(0.95) + (20)(0.10)(0.90) + (20)(0.02)(0.98) \\ &\quad + (20)(0.02)(0.92) + (20)(0.04)(0.96) = 5.382\end{aligned}$$

ومن ثمَّ ينتج لدينا أنَّ $\sigma = \sqrt{\mathbf{var}\,S_{100}} = 2.32$، ولننظر الآن إن كانت المتسلسلة $\sum_{k=1}^{\infty} p_k \cdot (1 - p_k)$ متباعدة أم لا؟

من أجل ذلك سنفترض جدلاً أنَّ هذه المتسلسلة العددية متقاربة، فعندئذ يجب أن يسعى حدّها العام إلى الصفر عندما $n$ يسعى إلى اللانهاية (وذلك بحسب الشرط اللازم لتقارب المتسلسلات العددية)، وهكذا ينتج لدينا أنَّ $\mathbf{min}\,\{p_n, 1 - p_n\} \xrightarrow[n \to \infty]{} 0$، وبناءً على ذلك فإنَّ المتتالية العددية $(p_n)_{n \in \mathbb{N}}$ إمَّا أن تحوي متتالية جزئية متقاربة من الصفر أو أن تحوي متتالية جزئية متقاربة من الواحد، وتفسير ذلك من أجل هذا المثال هو أنَّ المتسلسلة $\sum_{k=1}^{\infty} p_k \cdot (1 - p_k)$ تكون متقاربة إذا كانت معظم شحنات المصابيح التي قمنا باستلامها إما شحنة مصابيح سليمة تماماً أو معيبة تماماً، ولكن من خبرات السنوات الطويلة في إنّتاج المصابيح ومن عدد المصابيح المعيبة التي نحصل عليها في اختبارات الجودة اتضّح لنا أنَّه لا يمكن أن تحدث لدينا مثل هذه الحالة. إذن فرضنا الجدلي من أنَّ المتسلسلة $\sum_{k=1}^{\infty} p_k \cdot (1 - p_k)$ متقاربة غير صحيح، ومن أجل إزالة هذا التناقض يجب أن تكون هذه المتسلسلة متباعدة.

الآن، وبحسب التطبيق الأخير يكون لدينا $S_{100} \overset{\infty}{\sim} \mathbf{N}(94.2\,,\,2.32)$، أي إنَّ للمتغيِّر العشوائي $S_{100}$ توزيعًا طبيعيًّا $\mathbf{N}(94.2\,,\,2.32)$ على وجه التقريب، ومن ثمّ يمكننا أن نكتب الآتي:

$$P(S_{100} \le 89) = P\left(\frac{S_{100} - 94.2}{2.32} \le -2.25\right) = P(Z \le -2.25) \approx \Phi(-2.25) = 0.0122$$

علماً أنّنا قد قمنا باستعيّار المتغيّر العشوائي $S_{100}$ من خلال وضع $Z := \frac{S_{100} - 94.2}{2.32}$، فيكون لدينا $Z \sim \mathbf{N}(0\,,1)$. إنَّ قيمة الاحتمال الأخيرة تُعدّ صغيرة نسبياً، وتعني بدورها أنَّ قيم $p_k$ المُعطاة سابقاً أمر مبالغ فيه وينطوي على تفاؤل زائد.

### (٨،٣،٥) مسألة غنيدينكو Gnedenko Problem

قمنا فيما سبق بدراسة تقارب متتاليات دوال التوزيع لمجاميع متغيِّرات عشوائية مُستعْيرة (تمّ استعيارها) من دالّة التوزيع الطبيعي، ولكن في الحالة العامّة ليس بالضرورة أنَّ تقارب متتاليات دوال التوزيع لتلك المتغيِّرات العشوائية من التوزيع الطبيعي. هنا سوف لن نبحث في الإجابة العامّة على هذه المسألة، ولكن سنقدِّم دراسة لحالة خاصّة عندما يكون لدينا متتالية متغيِّرات عشوائية متقطِّعة ومستقلّة فقط. إنَّ هذه الحالة الخاصّة لدراسة التقارب تُدعى **نظرية النهاية المركزية الموضعية** Local Central Limit Theorem، وهذا يعني دراسة تقارب متتالية دوال التوزيع الاحتمالية من دالّة توزيع احتمالية ما يمكن تعيينها، وفي هذا المجال تكون أكثر الدراسات مركَّزة على المتغيِّرات العشوائية المتقطِّعة المستقلّة والمتطابقة التوزيع التي قيمها دوال خطية لأعداد صحيحة من الشكل $h \cdot k + a$ علماً أنَّ $a\ ,\ h \in \mathbb{R}$ مع $h > 0$ و $k \in \mathbb{Z}$ . إنَّ أمثال هذه التوزيعات تُدعى **توزيعات شبكية** Nets Distribution ، والعدد $h$ يُدعى **عرض خطوة التوزيع**. إنَّ العدد $h$ (وكما هو واضح) ليس إلاَّ الفرق لقيمتين من قيم المتغيِّر العشوائي $X_k$ ، وإذا كان $h$ هو أكبر قاسم مشترك لتلك الفروق فعندئذ $h$ يُدعى **عرض الخطوة الأعظمي**.

نقدم فيما يلي إحدى نظريات النهاية المركزية التي ينطبق عليها التمهيد السابق. إنَّ هذه النظرية تُعرف باسم نظرية غنيدينكو (وسنقبلها دون برهان) تعود إلى الرياضياتي الروسي **غنيدينكو** (1912-1995) Boris Vladimirovich Gnedenko الذي تتلمذ على يد الرياضياتي الروسي كلموغوراف.

### (٨،٣،٥،١) نظرية غنيدينكو Gnedenko Theorem

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية متقطِّعة فوق فضاء احتمالي $[\,\Omega,\mathscr{A},P\,]$ ، فإذا كانت هذه المتغيِّرات العشوائية تملك عزوماً ابتدائية من المرتبة الثانية، ومستقلّة، ومتطابقة في التوزيع، وتأخذ قيماً صحيحةً باحتمال موجب تماماً، فعندئذ من أجل كل $x$ من $\mathbb{R}$ تكون العلاقة الآتية مُحقَّقة **إذا وفقط إذا** كان عرض الخطوة الأعظمي لتوزيعات المتغيرات العشوائية $X_n$ يساوي الواحد من أجل كل $n \in \mathbb{N}$:

$$\lim_{n\to\infty}\left[\sigma\sqrt{n}\;\;\alpha_n(x)-\frac{1}{\sqrt{2\pi}}\,\mathbf{e}^{-\frac{x_{n,x}^{2}}{2}}\right]=0 \qquad [8,24]$$

علماً أنَّه من أجل كل $n \in \mathbb{N}$ و $x$ من $\mathbb{R}$ لدينا:

$$\alpha_n(x) := P\left(\sum_{i=1}^{n} X_i = x\right) \qquad \& \qquad x_{n,x} := \frac{x-n\cdot\mu}{\sigma\sqrt{n}}$$

$$\mu = \mathbf{E}X_n \qquad \& \qquad \sigma = +\sqrt{\mathbf{var}X_n} > 0$$

### (٨،٣،٥،٢) ملاحظات

١ – إذا كان عرض الخطوة الأعظمي يساوي الواحد فعندئذ يُقال عن المتغيِّر العشوائي إنَّه يحقِّق الشرط **[w]**.

٢ – يُلاحظ أنَّ نظرية النهاية المركزية الموضعية لـ موافير– لابلاس هي حالة خاصّة من النظرية السابقة، وذلك عندما يكون:

$$P(X_n = 1) = p \qquad \& \qquad P(X_n = 0) = 1-p \qquad ;\forall n \in \mathbb{N}$$

نقدم فيما يلي مبرهنة تعطينا إحدى النتائج المتعلِّقة بتحقُّق الشرط [8–24] .

### (٨،٣،٥،٣) مبرهنة

ليكن $X$ متغيِّراً عشوائياً فوق فضاء احتمالي $[\,\Omega,\mathscr{A},P\,]$ ، ويأخذ قيماً صحيحة فقط، ويُحقِّق الشرط **[w]**، فعندئذ تُحقِّق الدالّة

المميِّزة لهذا المتغيِّر العشوائي البندين الآتيين:

١- لدينا $\varphi_X(2\pi)=1$ .

٢- من أجل $0<|t|<2\pi$ لدينا $|\varphi_X(t)|<1$

البرهان: من أجل برهان البند:

١- سنفترض أنَّ $p_k = P(X=k)$ لكل $k$ من $\mathbb{Z}$، فعندئذ من أجل كل $t\in\mathbb{R}$ سيكون لـ $\varphi_X(t)$ العرض الآتي:

$$\varphi_X(t)=\sum_{k=-\infty}^{\infty} p_k \cdot \mathbf{e}^{\mathbf{i}k}$$

ومن ثَّم ينتج لدينا من أجل $t=2\pi$ ما يلي:

$$\varphi_X(2\pi)=\sum_{k=-\infty}^{\infty} p_k\, \mathbf{e}^{2\pi\,\mathbf{i}\,k}=\sum_{k=-\infty}^{\infty} p_k=1$$

وبذلك يتمّ إثبات الطلب الأول.

٢-سنفترض جدلاً أنَّه من أجل قيمة $t_o$ مع $0<|t_o|<2\pi$ لدينا العلاقة $|\varphi_X(t_o)|=1$ مُحقَّقة، فعندئذ يوجد عدد $a\in\mathbb{R}$ بحيث يكون $\varphi_X(t_o)=\mathbf{e}^{\mathbf{i}a}$، ومن ثَّم يمكننا أن نكتب العلاقة الآتية:

$$\int_{-\infty}^{+\infty} \mathbf{e}^{\mathbf{i}t_o x}\, dF_X(x)=\mathbf{e}^{\mathbf{i}\,a}$$

وبضرب طرفي العلاقة الأخيرة بـ $\mathbf{e}^{-\mathbf{i}a}$ ينتج لدينا ما يلي:

$$\int_{-\infty}^{+\infty} \mathbf{e}^{\mathbf{i}(t_o x-a)}\, dF_X(x)=1$$

ومن مُعطيات الفرض ($X$متقطّع بقيم صحيحة) يمكننا أن نكتب:

$$\int_{-\infty}^{+\infty} \cos(t_o x-a)\, dF_X(x)=\sum_{k=-\infty}^{+\infty} p_k \cos(t_o x-a)=1$$

و تكون هذه العلاقة الأخيرة مُحقَّقة فقط عندما يكون من أجل أيّة قيمة $k\in\mathbb{Z}$ لدينا العلاقة $\cos(t_o x-a)=1$ مُحقَّة، ومنه ينتج لدينا، أنَّه من أجل كل $k\in\mathbb{Z}$ يكون لـ $X$ القيم الآتية:

$$x_k=\frac{a}{t_o}+b_k\,\frac{2\pi}{t_o} \qquad ;\ b_k\in\mathbb{Z}$$

ولكن لدينا بالفرض $x_k$ هي قيم صحيحة، ولذلك سيكون للفرق بين أزواج القيم $x_k$ الشكل $s\cdot h$ علماً أنَّ $s$ هو عدد صحيح ما، وأمَّا $h=\dfrac{2\pi}{|t_o|}$ فهو أكبر قاسم مشترك لفروقات القيم $x_k$، وبسبب أنَّ $|t_o|<2\pi$ ينتج لدينا أنَّ $h>1$ وهذا يناقض الشرط [ $w$ ]، ولإزالة هذا التناقض يجب أنَّ يكون من أجل $0<|t|<2\pi$ لدينا $|\varphi(t)|<1$، وبهذا يتمُّ المطلوب.

**تطبيق (٤,٥,٣,٨)**

لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية متقطِّعة فوق فضاء احتمالي $[\Omega,\mathscr{A},P]$، وتملك عزوماً ابتدائية من المرتبة الثانية، مستقلّة، متطابقة في التوزيع وتأخذ قيماً صحيحةً لها الشكل الآتي:

$$x_{n,k} = a + h \cdot k \qquad ;\forall k \in \mathbb{Z}$$

مع $h > 0$ هو القاسم المشترك الأكبر لفروقات القيم بين الأزواج $x_{n,k}$. الآن لو أخذنا:

$$Z_n = \frac{X_n - a}{h} \qquad ;\forall n \in \mathbb{N}$$

فعندئذ المتغيِّرات العشوائية $X_n$ تأخذ قيماً صحيحة من أجل كل $n \in \mathbb{N}$، وكذلك تُحقِّق الشرط [ $w$ ] أيضاً، وهكذا يمكن تطبيق نظرية النهاية المركزية الموضعية لـ غنيدينكو على المتتالية $(X_n)_{n\in\mathbb{N}}$ وتكون العلاقة [8–24] مُحقَّقةً من أجلها.

## هذا ما تيسَّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل

# تمارين الفصل الثامن

١- تقدَّم 1000 شخص إلى اختبار مقابلة من أجل العمل بوظائف إدارية، فإذا علمت أنَّ احتمال رسوب أي شخص في هذا الاختبار يساوي 0.10 ، وأنَّ الاختبارات لهؤلاء الأشخاص تتمّ بشكل مستقل لكل منهم عن الآخر، فعندئذٍ:

أ- ما هو احتمال رسوب 50 شخصاً على الأقل؟

ب- ما هو احتمال أنَّ يكون عدد الأشخاص الذين سيرسبون في هذا الاختبار ما بين 50 و100 شخص؟

٢- تقوم طائرة مدنية برحلات ما بين المدن أو الدول، وبشكل مستقل كل منها عن الأخرى، فإذا كان $X_n$ يصف حالة السلامة لوصول الطائرة في الرحلة رقم $n$، وكان من أجل كل $n \in \mathbb{N}^o$ لدينا العلاقة الآتية مُحقَّقةً:

$$P\left(X_n = 1\right) = \frac{1}{2^n}$$

علماً أنَّ $X_o$ يمثِّل حالة الطائرة قبل الإقلاع الأول (أي إنَّها كانت سليمةً باحتمال يساوي الواحد)، وأنَّ الـ 1 يعني الوصول إلى هدفها دون مخاطر تُذْكر، ولنأخذ:

$$S_n = X_1 + X_2 + \ldots + X_n \qquad ; n \in \mathbb{N}^o$$

فعندئذٍ هل ستتقارب المتتالية $\left(\frac{S_n}{n}\right)_{n \in \mathbb{N}}$ بالتوزيع، وإلى أين ستتقارب؟

٣- في استطلاع للرأي حول خدمة البلدية في مدينة ما شمل 10000 شخص وبشكل مستقل كل منهم عن الآخر، وبحيث تكون الإجابة (راضٍ أو غير راضٍ)، فإذا افترضنا أنَّ أي الإجابتين لها الاحتمال نفسه، فما هو احتمال أن يكون عدد من الأشخاص الذين أجابوا بعبارة "راضٍ" ما بين 4985 و5005؟

٤- لتكن $\left(X_n\right)_{n \in \mathbb{N}}$ متتالية متغيِّرات عشوائية مستقلّة ومتطابقة التوزيع مع توزيع متغيِّر عشوائي $X$ خاضع للتوزيع الهندسي بمَعْلَمة $p = 0.25$، ولنأخذ $S_n = \sum_{i=1}^{n} X_i$ من أجل كل $n \in \mathbb{N}$، والمطلوب حساب الاحتمال $P\left(70 < S_{300} < 130\right)$.

٥- نقوم برمي حجر نرد متوازن لعدد كبير جداً من المرَّات، وليكن $X_n$ متغيِّراً عشوائياً راصداً لنتيجة الرمية ذات الرقم $n$ مع $n \in \mathbb{N}$، والمطلوب حساب الاحتمالين الآتيين:

$$P\left(\bar{X}_{1000} < 3\right) = ? \qquad \& \qquad P\left(\bar{X}_{1000} > 5\right) = ?$$

ماذا تلاحظ، وما السبب في ذلك، علماً أنَّ $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$.

٦- تقوم دارة إلكترونية بتوليد عشوائي لأعداد حقيقية $0 \leq x \leq 1$ وبشكل مستقل كل منها عن الآخر، فإذا علمت أنَّ التوزيع المنتظم المستمر على الفترة $[0, 1]$ هو المعتمد في عشوائية القيم المولَّدة، فعندئذ:

أ- إذا ولَّدت هذه الدارة 100 رقمًا عشوائيًّا فما هو احتمال أن يكون لمتوسط هذه القيم قيمة أكبر من 0.505، وكذلك ما هو احتمال أن يكون لمتوسط هذه القيم قيمة أصغر من 0.495؟

ب- إذا ولَّدت الدارة 1000 رقم عشوائي فما هو احتمال أن يكون لمتوسط هذه القيم قيمة أكبر من 0.505، وكذلك ما هو احتمال أن يكون لمتوسط هذه القيم قيمة أصغر من 0.495، وماذا تلاحظ؟

٧- لدينا سبع جامعات لكل منها نظام ومنهج خاص في التعليم، وفي كل جامعة عددٌ ما من الطلاب المتعثرين، فإذا كان $p_k$ هو احتمال الحصول على طالب متعثر من الجامعة رقم $k$، فعندئذ بفرض أنَّ:

$$p_1 = 0.95 \quad \& \quad p_2 = 0.93 \quad \& \quad p_3 = 0.98$$

$$p_4 = 0.99 \quad \& \quad p_5 = 0.92 \quad \& \quad p_6 = 0.96 \quad \& \quad p_7 = 0.98$$

وقمنا بسحب عشوائي لخمسين طالباً من كل جامعة، ومن ثمَّ الكشف عن وضع جميع الطلاب الذين تمَّ اختيارهم فوجدنا من بينهم 50 طالباً متعثِّراً، فهل تثير لديك هذه النتيجة الشك تجاه القيم $p_k$ المقدَّمة آنفًا، ولماذا؟

٨- لتكن $(X_n)_{n\in\mathbb{N}}$ متتالية متغيِّرات عشوائية مستقلّة ومتطابقة في التوزيع مع متغيِّر عشوائي $X$ خاضع لتوزيع كوشي المعياري، والمطلوب تعيين توزيع المتوسط $\overline{X}_n$ (استفد من نتيجة التمرين / ١٦/ من تمارين الفصل السابع).

٩- نقوم بإجراء تجارب عشوائيَّة مستقلّة ونرصد من خلالها تحقُّق حادث ما $A$ ، فعندئذكم مرَّة يجب أنَّ نكرِّر هذه التجارب حتى يكون للفرق بين احتمال الحادث $A$ وتكراره النسبي أقل من 0.01 باحتمال قدره 0.99 على الأقل؟

١٠- لدينا مصنع ينتج دارات متكاملة من نوع معيَّن بحيث إن نسبة الدارات المعطَّلة في إنَّتاجه تساوي 0.1%، قمنا بشراء 10000 دارة من إنَّتاج هذا المصنع، فما هو احتمال أنَّ يكون عدد الدارات المعطَّلة لدينا يقع ما بين 400 و600 دارة؟

١١- يوجد في مستودع لصناعة الأواني الخزفية أنواع مختلفة من صحون الطعام معبأة في صناديق بحيث يحوي كل صندوق عدداً كبيراً جداً من الصحون (ولكن منته). نأخذ عشوائيًّا أربعة صناديق $B_1$ ، $B_2$ ، $B_3$ و $B_4$ من مجمل الإنتاج الموجود في المستودع، ومن ثمَّ نقوم بسحب عشوائي لخمسة وعشرين صحناً من كل صندوق، وبعد ذلك نتفحَّص هذه الصحون فوجدنا منها 10 صحون معيبة، فإذا علمنا أنَّ احتمال الحصول على صحن معيب (لا يحقِّق المواصفات المطلوبة) من هذه الصناديق الأربعة هو:

$$p_1 = 0.05 \quad \& \quad p_2 = 0.02 \quad \& \quad p_3 = 0.15 \quad \& \quad p_4 = 0.10$$

على الترتيب، وأنَّه من خبرتنا السّابقة في هذا المجال من غير الممكن أن يكون كامل المنتج سليماً أو كلّه معيباً، فهل تثير هذه النتيجة في نفسك الشك تجاه افتراضنا للقيم $p_k$ المُعطاة (علّل ذلك)؟

# الفصل التاسع

## مدخل إلى نظرية المعاينة

## AN INTRODUCTION TO SAMPLING THEORY

### (٩,١) مفاهيم ومصطلحات أساسية

**Basic concepts and expressions**

سنذكِّر أولاً ببعض المفاهيم التي تخص هذا الموضوع وقدِّمت في الفصل الأول من هذا الكتاب. لقد عرضنا في ذلك الفصل تمهيداً مبسطاً حول علم الإحصاء وأهميته في مختلف العلوم، ومن ثمَّ تطرقنا إلى مفهوم المجتمع الإحصائي وحجمه، حيث ذكرنا أنَّ المجتمع الإحصائي هو أي تجمُّع لأشياء تجمَع بينها صفة مشتركة واحدة على الأقل لتكون محل دراسة لهدف معيَّن، ومكوّنات هذا المجتمع تُدعى أفراداً أو عناصر، وأمّا عدد العناصر المكوِّنة للمجتمع الإحصائي فإنَّها تُدعى حجم المجتمع، وفي حال كان حجم المجتمع منتهياً يُرمز لحجمه بحرف لاتيني كبير من قبيل $N$ أو $M$ أو.... بعد ذلك أوضحنا ضرورة التعامل مع العيّنة للحصول على البيانات التي نودّ استخدامها في دراسة إحصائية ما، وذكرنا أنَّ العيّنة هي جزء يسحب من المجتمع الإحصائي لتنفيذ الدراسة الإحصائية المطلوبة متجنبين بذلك استخدام جميع عناصر المجتمع (المسح الشامل) لأسباب معيَّنة تفرضها طبيعة الدراسة بحيث إمّا أن نكون غير قادرين على التعامل مع كل عناصر المجتمع، أو أنَّ ذلك يجعل الدراسة عقيمةً وغير ذات معنى، أو أنَّ التعامل مع كل عناصر المجتمع مكلف مادياً أو زمنياً (أو هنا لا يفيد الحصر). كذلك ذكرنا أنَّ العلم الذي يبحث في كيفية أخذ العيّنات يُعرف باسم **نظرية المعاينة**، وأوضحنا أنَّ هذه النظرية تُعدُّ جزءاً لا يتجزأ من العلوم الإحصائية المختلفة، وأنَّ تطبيقها في البحوث العلميّة تمكننا من الحصول على نتائج سريعة وبكلفة زهيدة نسبياً إذا ما قورنت بكلفة المسح الشامل، وأكثر من ذلك ينظر في كثير من الحالات إلى هذه النظرية على أنَّها الوسيلة الوحيدة المتاحة في البحث العلمي.

الآن، وبعد هذه المقدِّمة الموجزة سنقوم باستكمال بعض المصطلحات والمفاهيم المستخدمة في إطار نظرية المعاينة، ولكن بشكل موجز تقريباً حيث يمكن لمن يود الاطلاع على المزيد الرجوع إلى المراجع ذات الصلة التي ومنها [17] و[18] و[24] و[33] و[40] و[45] و[64] في فهرس مراجع هذا الكتاب.

#### (٩, ١, ١) تعريف (الإطار Framework)

الإطار هو مجموعة كل الرموز أو الأرقام أو الأحرف التي تُمثِّل عناصر المجتمع الإحصائي عند إعداده من أجل سحب عيّنة لإخضاعها للدراسة.

#### (٩, ١, ٢) تعريف (وِحْدَة المعاينة Sampling Unit)

وِحْدَة المعاينة هي مصطلح يُطلق على مجموعة العناصر من المجتمع الإحصائي التي تتميَّز عن العناصر الأخرى بصفة معيّنة.

في الواقع هذه الصفة التي تُميِّز العناصر تُشكِّل البنية الأساسيَّة لعملية أخذ جزء أو أكثر من المجتمع الذي هو قيد الدرس، كأنَّ نأخذ على سبيل المثال من مجتمع الطلبة في بلد ما جزءاً من الطلبة ذوي التقدير الممتاز من أجل معرفة متوسط الدرجة لمحقِّقي هذا التقدير، و من ثَمَّ وحْدَة المعاينة هنا تتمثل بالطلبة ذوي التقدير الممتاز، وهكذا نلاحظ أنَّ وحْدَة المعاينة التي تصلح لنوع مُعيَّن من الدراسة قد لا تصلح لنوع آخر من الدراسات الأخرى المتعلّقة بالمجتمع الإحصائي نفسه.

في الحقيقة إنَّ عمليّة سحب العيّنات تتطلب تمعّناً وتفكيراً مليّاً في آليتها، وذلك لأنَّ نتائج السحب يبنى عليها قرارات مصيرية بخصوص الدراسة الإحصائية، فكلما كانت طريقة السحب أكثر ملاءمة للدراسة الإحصائية أدى ذلك إلى نتائج أفضل.

### (٩,١,٣) طرائق سحب العيّنة Methods of Sampling:

في الواقع يمكن إجراء عمليّة سحب العيّنة من المجتمع بإحدى طريقتين:

#### أ- السحب مع الإعادة (أو الإرجاع) Draw with Replacement:

في هذا النوع من السحب يُعاد العنصر المسحوب (من المجتمع) إلى المجتمع ثانية بعد تطبيق الدراسة الإحصائية عليه، ومن ثَم يمكن للعنصر أنَّ يظهر عدداً كيفياً من المرّات خلال عمليّات السحب، فعلى سبيل المثال لو افترضنا أنّه لدينا وعاء يحوي $N$ كرة (المجتمع حجمه $N$) متماثلة ومرقمة من 1 وحتى $N$، وقمنا بخلط هذه الكرات جيداً، ومن ثمَّ سحب كرة من الوعاء وتسجيل رقمها، ثمَّ نعيدها إلى الوعاء ثانيةً، ونخلط الكرات من جديد بشكل جيد، وبعد ذلك نسحب كرة ثانية ونسجل رقمها ثمَّ نعيدها إلى الوعاء من جديد، وهكذا... نكرِّر هذه العمليّة لـ $n$ مرّة متتالية، فعندئذ نكون قد حصلنا على عيّنة حجمها $n$، فعندئذ يُقال عن هذه العيّنة إنَّها ناتجة عن السحب مع الإعادة (أو الإرجاع، أو الإحلال). لاحظ هنا أنّه يمكننا أن نختار عناصر هذه العيّنة من المجتمع بعدد من الطرائق يساوي $N \cdot N \cdot \ldots \cdot N = N^n$ طريقة.

#### ب- السحب بدون إعادة (أو بدون إرجاع) Draw without Replacement:

في هذا النوع من السحب لا يُعاد العنصر المسحوب إلى المجتمع ثانية بعد تطبيق الدراسة الإحصائية عليه، وإنَّما يُحَيَّد جانباً (إمّا لعدم الرغبة في استخدامه مرّة ثانية أو بسبب تلفه نتيجة للبحث والدراسة)، ومن ثمَّ فإنَّه من غير الممكن ظهور العنصر أكثر من مرّة واحدة خلال عمليّات السحب، فعلى سبيل المثال لو افترضنا أنّه لدينا وعاء يحوي $N$ كبسولة عقار دوائي من إنتاج مصنع للدواء (ولتكن على سبيل المثال كبسولات سيتامول الخافض للحرارة)، وقمنا بخلط هذه الكبسولات جيداً، ومن ثمَّ نسحب إحداها ونخضعها للتحليل المخبري للتحقُّق من نسبة المادة الدوائية الفعَّالة فيها، فعندئذ نلاحظ أنّه لا يمكننا إعادة هذه الكبسولة إلى الوعاء الحاوي للكبسولات؛ لأنَّها قد أتلفت، وبعد ذلك نخلط الكبسولات المتبقية من جديد بشكل جيد، ونقوم بسحب كبسولة ثانية وتنفيذ الدراسة الإحصائية عليها، وهكذا نستمر في تكرار هذه العمليّة لـ $n$ مرّة متتالية، فنكون بذلك قد حصلنا على عيّنة بحجم $n$، وفي هذه الحالة يُقال عن العيّنة إنَّها ناتجة عن السحب بدون إعادة (أو إرجاع، أو إحلال).

نلاحظ هنا أنّه لا يمكن للعنصر أنَّ يظهر أكثر من مرّة واحدة خلال عمليّات السحب، ولهذا السبب فإنَّه لدى التعامل مع هذا النوع من السحب للعيِّنات يجب النظر إلى حجم المجتمع والعيِّنة بحيث لا يؤثِّر سحب العيّنة في حجم المجتمع، ومن ثمَّ يفترض أنَّ يكون حجم المجتمع كبيراً بقدر كاف إذ ما قُورن بحجم العيّنة حيث ذكرنا سابقاً أنّه في مثل هذه الحالات من المقبول إحصائيّاً أن يكون $N \ge 20n$. كذلك يُلاحظ أنّه يمكن اختيار عناصر هذه العيّنة من المجتمع الإحصائي بعدد من الطرائق يساوي:

$$N \cdot (N-1) \cdot (N-2) \cdot \ldots \cdot (N-n+1) = \frac{N!}{(N-n)!}$$

## (٩,٢) تصنيف العينات

### Classification of Samples

مما تقدّم يتبيّن لنا أنَّ استخدام مفهوم العيّنة في الدراسة (أو البحث الجزئي) إنّما هو وسيلة لتعميم ما تمَّ التوصل إليه من نتائج على المجتمع بأكمله، على اعتقادنا أنَّ هذه العيّنة المأخوذة من المجتمع هي مُمثِّل مقبول للمجتمع، وذلك وفقاً لمعايير نعدّها مقنعة لأجل العمل الذي نقوم به، وبحيث لا يؤدي ذلك إلى تناقضات علميّة، وفي هذا الإطار تبرز الوسائط العامّة ومهارة الباحث وحدسه في تعيين نوع العيّنة التي يجب التعامل معها، فإمّا أن تكون العيّنةً نتيجة لاختيارات عناصر من المجتمع دون تدخل من قبل الباحث، ومن ثمَّ يسيطر على هذا النوع من العيّنات طابع العشوائية في سحب العناصر من المجتمع، أو أن تكون نتيجة لاختيارات مُحدَّدة من عناصر المجتمع بتدخل مباشر من قبل الباحث، وذلك اعتقاداً منه أنَّ هذه العناصر المنتقاة هي مُمثِّل جيد للمجتمع المدروس، ومن ثمَّ العيّنة الناتجة ستكون ذات طابع عمدي أو قصدي (ليست عشوائية).

إذاً ، وبناءً على ما ذُكر يمكن تصنيف العيّنات التي ستسحب من المجتمع الإحصائي إلى نوعين:

**الأول:** عيِّنات عشَوائية Random Samples ، وتكون عناصرها قد سُحبت بطرائق عشوائيَّة.

**الثاني:** عيِّنات عمديَّة (أو قصديَّة) Intentional Samples، وعناصرها تَنتج عن عمليّات انتقاء وفقاً للرأي الشخصي للباحث.

### (٩ , ٢ , ١) العيّنات العشوائية Random Samples

فيما يلي سنقوم بعرض موجز لبعض نماذج العيّنات العشوائية. علماً أنَّ العيّنة العشوائية هي عيّنة تسحب من المجتمع بطرائق عشوائية، وبما أنَّ الطرائق العشوائية كثيرة فإنَّه سيكون لدينا نماذج عديدة من العيّنات العشوائية. هذا من جانب، ومن جانب آخر، فإنَّ العيّنات العشوائية على مختلف أنواعها تجعل الاستعانة بنظرية الاحتمالات أمراً ممكناً في تعميم النتائج من العيّنة على المجتمع، ومن ثمَّ تقديم الاختبارات المتعلّقة بفرضيات حول هذا المجتمع، وكذلك في حساب الأخطاء المرتكبة في هذا التعميم.

في الحقيقة توجد طرائق كثيرة لسحب العيّنات العشوائية، ومعظمها معروف ومألوف لدى الكثير من النّاس مثل طريقة الكيس أو الصندوق، حيث يوضع بداخله أشياء (مثل قصاصات ورق دوِّن عليها أرقاماً أو أسماءً أو ...) تُمثِّل عناصر المجتمع، وبعضها الآخر خاص بالباحثين العلميين مثل طريقة جداول الأرقام العشوائية، وطريقة مونتكارلو، والمولّدات الإلكترونية للأرقام العشوائية (مولّدات أرقام عشوائية وفقاً لتوزيع احتمالي ما) الموجودة في الحاسبات الإلكترونية، ولكنّنا سنعرض هنا طريقة استخدام جدول الأرقام العشوائية فقط (حيث يمكن لمن يود الاطلاع على المزيد الرجوع إلى كتب نظرية المعاينة).

في معظم كتب الإحصاء توجد جداول إحصائية متنوِّعة، منها جداول تحوي خانات ذات عدد متساوٍ من الأرقام، فمنها مكوَّنة من خانة واحدة تشتمل على الأرقام من 1 وحتى 9 (ومن الممكن أن يبدأ الترقيم من الـ 0أيضاً)، أو مكوَّنة خانتَين تشتمل على الأرقام من 01 وحتى 99، أو مكوَّنة من ثلاث خانات تشتمل على الأرقام من 001 وحتى 999، أو أكثر من ذلك، وتتوارد هـذه الأرقام بشكل عشوائي (مُولَّدة باستخدام بعض التوزيعات الاحتمالية، وغالباً التوزيع المنتظم)، وتعرف هذه الجداول باسم جداول الأرقام العشوائية Table of Random Digits حيث تمَّ نشر أول جدول للأرقام العشوائية من قبل الإحصائي الإنكليزي لونارد هنري تيبت Leonard Henry Caleb Tippett (1902-1985) وذلك في عام 1927. إنَّ هذه الجداول على أنماط وأشكال مختلفة كما ذكرنا ذلك آنفاً، ومنها على سبيل المثال النماذج الآتية:

| | | | |
|---|---|---|---|
| 6 1 5 3 | 5 9 8 1 | 4 8 7 8 | 9 2 8 0 |
| 4 1 5 9 | 1 9 3 4 | 2 9 1 0 | 0 2 0 4 |
| 6 7 5 0 | 5 0 8 5 | 4 8 7 8 | 1 2 8 5 |
| 5 0 5 0 | 7 9 8 9 | 4 8 7 8 | 9 7 8 0 |

جزء من جدول أحادي الخانات

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 65 | 10 | 51 | 33 | 95 | 99 | 48 | 11 | 14 | 08 | 17 | 78 |
| 41 | 51 | 75 | 99 | 91 | 19 | 33 | 84 | 02 | 69 | 71 | 20 |
| 26 | 07 | 05 | 50 | 35 | 01 | 78 | 25 | 74 | 88 | 47 | 28 |
| 05 | 70 | 71 | 09 | 37 | 91 | 18 | 90 | 41 | 80 | 79 | 81 |
| 05 | 40 | 55 | 63 | 25 | 19 | 58 | 15 | 64 | 09 | 11 | 18 |
| 11 | 81 | 65 | 90 | 41 | 14 | 35 | 74 | 12 | 09 | 41 | 23 |
| 20 | 02 | 55 | 10 | 25 | 71 | 70 | 75 | 34 | 81 | 07 | 08 |
| 45 | 40 | 41 | 59 | 07 | 01 | 98 | 09 | 81 | 10 | 59 | 99 |

جزء من جدول ثنائي الخانات

| | | | |
|---|---|---|---|
| 67835 | 63314 | 50162 | 36738 |
| 25014 | 15460 | 11730 | 76548 |
| 05393 | 96770 | 56087 | 40875 |
| 13351 | 93994 | 22375 | 00953 |
| 85423 | 60698 | 01898 | 61414 |

جزء من جدول خماسي الخانات

| | | | | |
|---|---|---|---|---|
| 0310650 | 1098218 | 5451034 | 3363111 | 9125545 |
| 9957519 | 1563948 | 1001101 | 0904014 | 7521708 |
| 3420217 | 7991198 | 1054641 | 5030981 | 0110075 |
| 7134299 | 9012191 | 1777759 | 3210333 | 8354474 |
| 2085202 | 6456419 | 7117771 | 2065550 | 0012190 |
| 1755559 | 8870333 | 8354621 | 1105393 | 4496770 |
| 1250145 | 7154606 | 1117300 | 7765488 | 9614141 |

جزء من جدول سباعي الخانات

من أجل استخدام جدول الأرقام العشوائية نقوم أولاً بترقيم عناصر المجتمع بأرقام متسلسلة بدءاً من الواحد وصاعداً، ثمَّ نأخذ خانة من خانات هذا الجدول بطريقة عشوائية (كأن نغمض عينينا ونضع إصبعنا على الجدول، فنختار الخانة أو المجموعة لتلك الأرقام التي تحت الإصبع)، وبعد ذلك نسحب من المجتمع تلك العناصر التي يوافق رقمها الأرقام التي حصلنا عليها من الخانة المذكورة، حيث يجب الأخذ بالحسبان الأمور الآتية:

١- إذا كان الرقم الذي وقع عليه الاختيار يقع خارج نطاق الأرقام المستخدمة في ترقيم عناصر المجتمع، فإنَّه يُعدُّ لاغياً، ونتابع السـحب باستخدام الرقم الذي يليه.

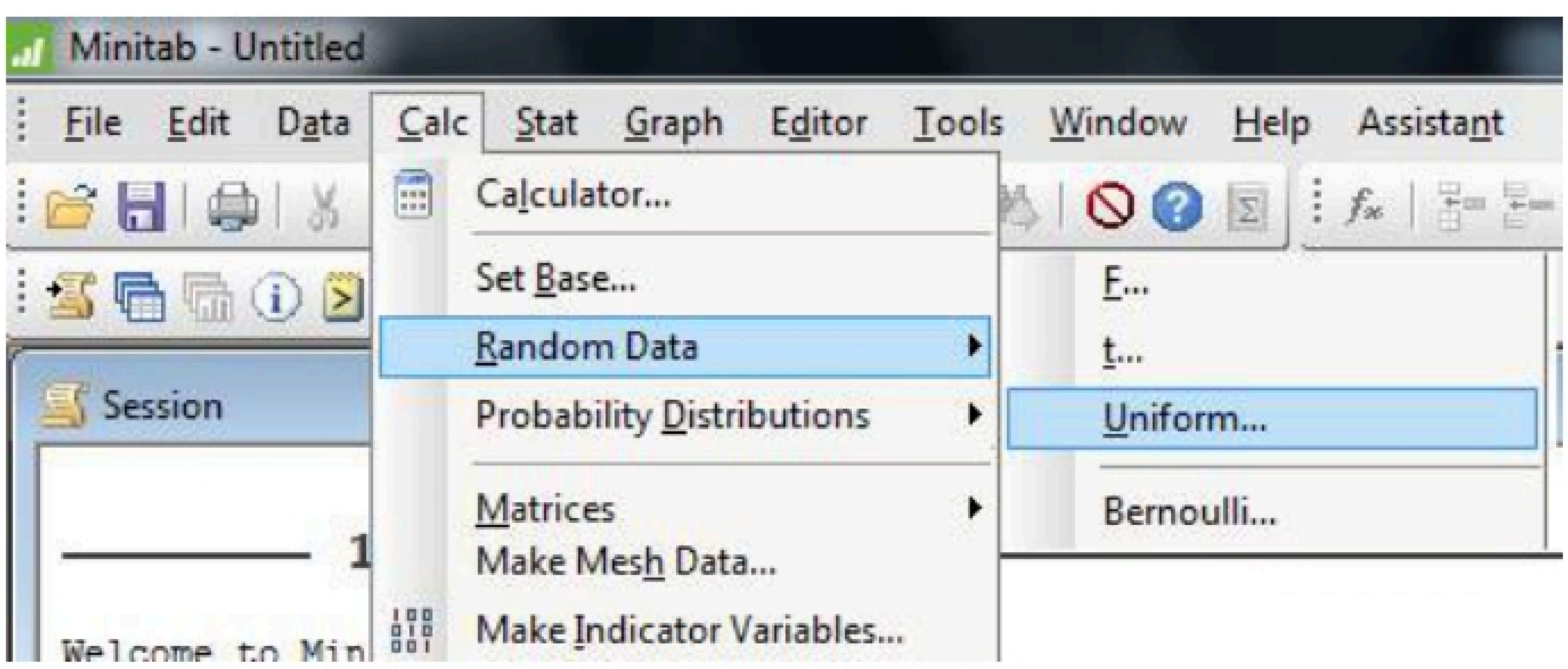
٢- إذا كنّا مهتمين بالسحب دون الإعادة فإنَّ الرقم الذي وقع عليه الاختيار يُعدُّ لاغياً إذا كان هذا الرقم هو تكرار لرقم سابق تمَّ سحبه، فعلى سبيل المثال لو كان لدينا مجتمعاً إحصائياً حجمه $N = 132$ وأردنا سحب عيّنة عشوائية بدون إعادة باستخدام جدول الأرقام العشوائية وعلى أن يكون حجمها يساوي $n = 10$، فعندئذ نختار جدولاً تكون فيه الأرقام بثلاث خانات على الأقل، وليكن النموذج الآتي الذي يحتوي أرقاماً بخمس خانات:

| | | |
|---|---|---|
| 56087 40875 13351 93994<br>22375 00953 22375 00953 | 39941 21225 93629 41469<br>16812 81542 04231 40875 | 93994 36009 76548 05393<br>67701 85423 60698 01898 |
| 00619 85423 60698 01898<br>61414 83525 76548 05393<br>96770 22909 29563 44018<br>64732 93589 61098 04393<br>48245 56087 40875 13351 | 41383 31555 12639 82667<br>74624 36348 77319 73408<br>58993 99410 48245 56087<br>22375 00953 22375 00953<br>04231 40875 13351 93994 | 64109 09497 76235 16210<br>89717 65997 02760 24359<br>36009 76548 05393 67705<br>71565 33413 81652 45554<br>56087 40875 13351 93994 |
| 52975 58993 99410 48245<br>56087 41383 31555 12639 | 10081 04231 40875 13351<br>93994 73152 14511 85285 | 86576 86944 93296 66456<br>47679 66810 41045 82830 |

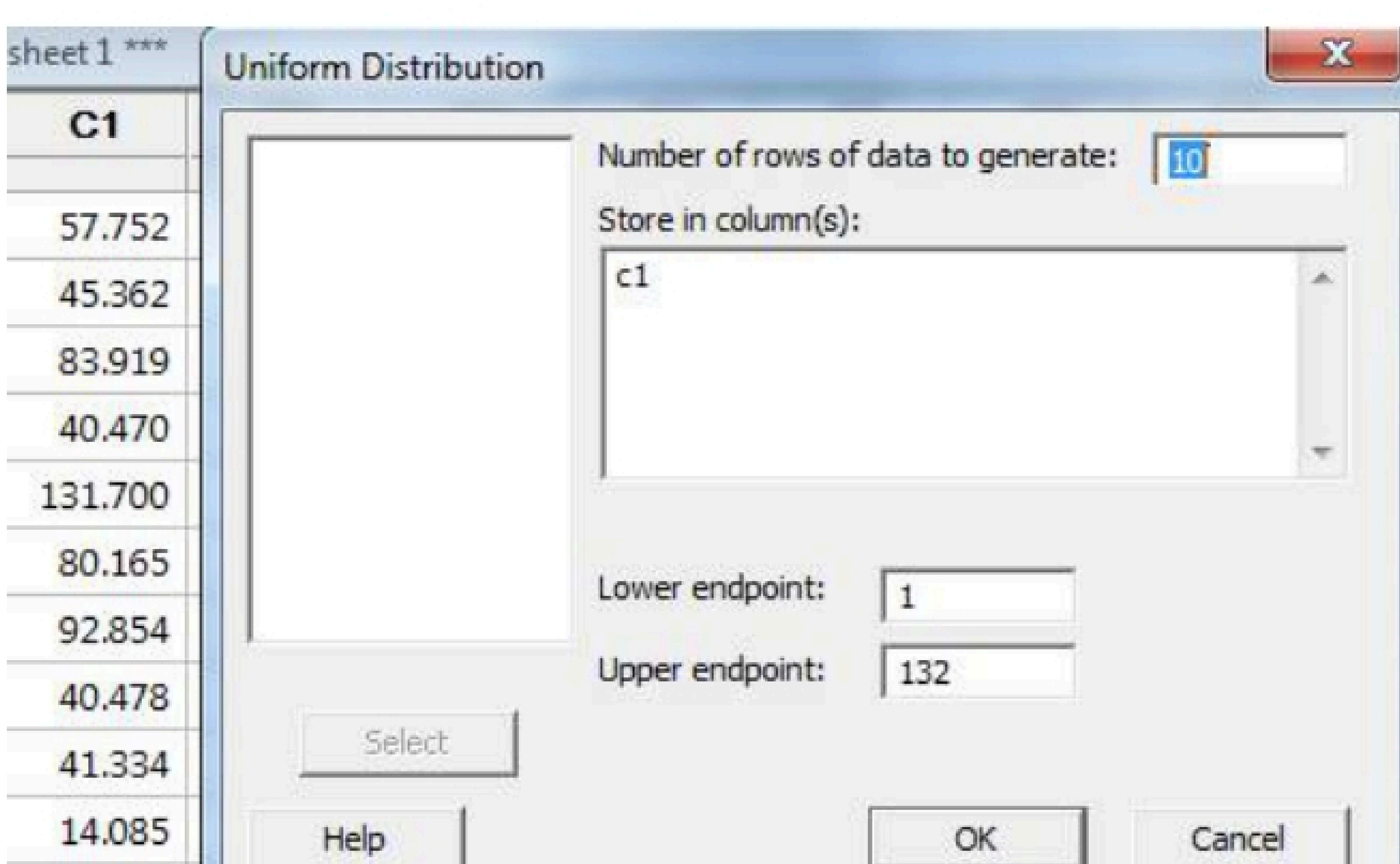
ثمَّ لنختر عشوائياً أحد هذه المجموعات، ولتكن على سبيل المثال المجموعة المحاطة بالمنحنى المغلق، فعندئذ أول عنصر يجب سحبه يحمل الرقم 41 (لأنَّ الرقم 413 غير مستخدم في الإطار)، والعنصر الثاني رقمه 31، والثالث 126، وعلى هذا النحو نتابع السحب فتكون أرقام العناصر التي يجب أن تسحب من المجتمع موافقة للأرقام الآتية:

**41 31 126 82 74 36 77 73 58 99**

كما يمكننا استخدام بعض البرامج الإحصائية لتوليد القيم العشوائية لسحب عيّنة عشوائية ومنها على سبيل المثال برنامج Minitab وذلك من خلال استخدام الخطوات الموضّحة في العرض الآتي:




فنحصل على اللوحة التالية التي نحدّد فيها حجم المجتمع والعينة والعمود الذي ستخزن عليه الأرقام المولَّدة.




فنحصل على القيم المعروضة في القسم الأيسر التي تُقرَّب إلى أقرب عدد صحيح (لأنَّ القيم المستخدمة في الإطار هي قيم صحيحة) فنجد العناصر التي يجب سحبها من المجتمع هي تلك العناصر التي تحمل الأرقام الآتية:

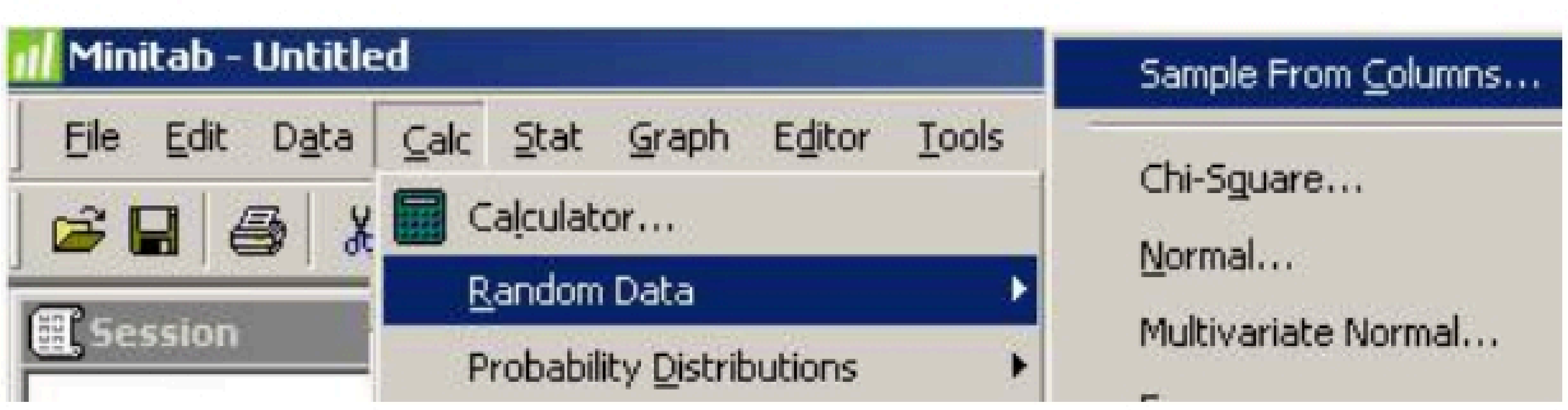
**58 45 83 40 132 80 93 40 41 14**

سنقوم فيما يلي بتقديم بعض أهم أنواع العيّنات العشوائية، ونبدأها بالعيِّنة العشوائية البسيطة.

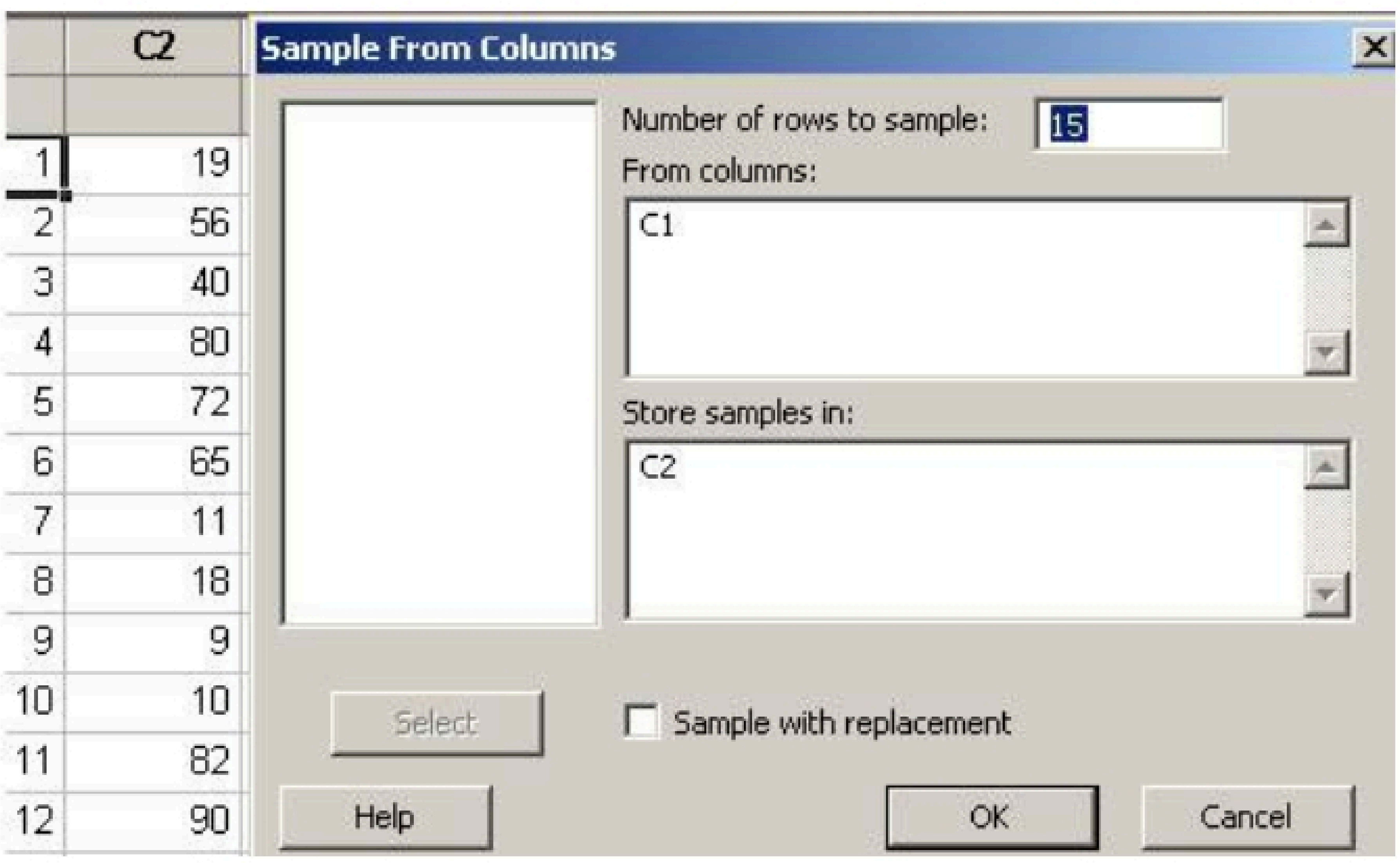
**(١, ١, ٢, ٩) العيّنة العشوائية البسيطة Simple Random Sample**

يتميَّز هذا النوع من العيّنات العشوائية بأنَّ لجميع عناصرها النصيب نفسه في الانتقاء مع أية عيّنة عشوائية أخرى لها الحجم نفسه وممكنة التشكيل من المجتمع نفسه، وهذا يوافق الحالة التي يكون فيها جميع عناصر المجتمع مستقلّة بعضها عن بعض ولها النصيب نفسه في الظهور، ولتوضيح ما سبق لنأخذ طلبة جامعة ما **X** كمجتمع إحصائي، ولنفترض أنّنا بصدد تعيين متوسط طول الطلبة في هذه الجامعة، وأنَّ جميع الطلبة لهم النصيب نفسه في الاختيار، وكما هو معلوم طول أي طالب لا يتأثّر بطول الطلبة الآخرين فرادى كانوا أم جماعات، فعندئذ إذا قمنا بسحب عشوائي لـ $n$ طالب من طلبة هذه الجامعة، فإنَّ هذه العيّنة المسحوبة هي عيّنة عشوائية بسيطة؛ وذلك لأنَّ أيّة عيّنة عشوائية أخرى ستسحب من هذا المجتمع ولها الحجم نفسه سيكون لها الاحتمال نفسه في انتقاء العيّنة السّابقة.

كما يمكن استخدام برنامج Minitab لتوليد عينات عشوائية بسيطة وفقاً للطريقة المقدمة سابقاً أو من خلال تدوين أرقام العناصر في أحد الأعمدة ومن ثمَّ طلب توليد عيّنة عشوائية من البيانات المدونة بالحجم المطلوب، فلو افترضنا أنَّه لدينا مجتمع حجمه $N=100$ وأردنا سحب عيّنة عشوائية بسيطة منه بحجم $n=10$ ، فعندئذ نحصل على أرقام عناصر هذه العيّنة باستخدام الأمر Calc من شريط الأوامر ومن ثمَّ الانتقال إلى Random Data وبعد ذلك اختيار Sample Form Column ومن ثمَّ اتباع الخطوات كما في العرضين الآتيين.




ومن ثمَّ




فنحصل نتيجة لذلك أنَّ عناصر العيّنة العشوائية البسيطة التي تحمل الأرقام الآتية:

| | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 19 | 56 | 40 | 80 | 72 | 65 | 11 | 18 | 9 | 10 | 82 | 90 | 37 | 71 | 39 |

### (٢, ١, ٢, ٩) العيّنة الرتيبة (أو المنتظمة) Systematic Sample

إنَّ هذا النوع من العيّنات العشوائية يقوم على الأسس الآتية:

١- إعطاء جميع أفراد المجتمع أرقاماً متسلسلة بدءاً من الواحد.

٢- تحديد قيمة عددية $s$ تمثّل طول فترة تُدعى فترة السحب ومُعطاة من خلال العلاقة الآتية:

$$s := \left\lceil \frac{N}{n} \right\rceil \qquad [9,1]$$

علماً أنَّ $N$ و $n$ يمثّل حجم المجتمع والعيّنة على الترتيب، وأمّا $\lceil x \rceil$ فهو أصغر عدد صحيح أكبر من $x$ حيث يشترط هنا أنَّ يكون حجم المجتمع محدوداً وذلك لأنَّه عندما يكون حجم المجتمع غير محدود فإنَّ العلاقة الأخيرة تفقد معناها.

٣- سحب عشوائي لعنصر من المجتمع يكون رقمه $u$ واقعاً بين الواحد و $s$ (أي إنّ $1 \le u \le s$ وذلك وفقاً لإحدى الطرائق العشوائية المتوفرة لدينا)، و هذا العنصر يُدعى عنصر البدء Initial Element.

٤- أخذ بقية عناصر العيّنة من عناصر المجتمع التي رقمها يوافق العدد $u + k \cdot s$ مع $k \in \mathbb{N}_{n-1}$ .

لتوضيح ذلك لنفترض أنَّه طلب منّا حساب متوسط عدد الكلمات في كلّ صفحة من كتاب يحوي 420 صفحة مرقمة بـ 1 وحتى 420، وذلك باستخدام طريقة المعاينة الرتيبة على أن يكون حجم العيّنة $n = 19$ . من أجل سحب مثل هذه العيّنة العشوائية نقوم أولاً بتعيين فترة السحب $s$ فنجدها تساوي:

$$s = \left\lceil \frac{420}{19} \right\rceil = \lceil 22.105 \rceil = 23$$

وبعد ذلك نسحب وبشكل عشوائي رقماً يقع بين 1 والـ 23، وليكن على سبيل المثال الرقم 12، فعندئذ الصفحة ذات الرقم 12 هي عنصر البدء في العيّنة (صفحة البدء في العيّنة)، ومن ثمَّ تكون الصفحات ذات الرقم $12 + 23k$ مع $k \in \mathbb{N}_{10}$ هي:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **12** | **35** | **58** | **81** | **104** | **127** | **150** | **173** | **196** | **219** |
| **242** | **265** | **288** | **311** | **334** | **357** | **380** | **403** | **426** | |

فنكون ذلك قد حصلنا على العيّنة الرتيبة التي سنستخدمها لتقديم متوسط عدد الكلمات في كل صفحة من صفحات الكتاب، حيث نقوم بعدّ الكلمات في كل صفحة من هذه الصفحات التابعة للعينة، ومن ثمَّ حساب المتوسط للقيم الأخيرة فتكون هي القيمة الافتراضية لمتوسط عدد الكلمات في كل صفحة من صفحات هذا الكتاب.

### (٣, ١, ٢, ٩) العيّنة الطبقية Stratified Sample

إنَّ هذا النوع من العيّنات العشوائية يقوم على الأسس الآتية:

١- تقسيم المجتمع إلى فئات بحيث تكون عناصر هذه الفئات متجانسة قدر المستطاع، وعندئذ تُدعى كل فئة من هذه الفئات طبقة.

٢- سحب عيّنة عشوائية بسيطة من كل طبقة من هذه الطبقات حيث يمكن النظر إلى كل طبقة من هذه الطبقات على أنَّها مجتمع جزئي متجانس تقريباً، وأمّا حجم هذه العيّنات البسيطة فيُحدَّد وفقاً لمعايير مختلفة من أهمها:

أ- أن يكون وفقاً لنسبة مئوية مُحدَّدة من تعداد عناصر كل طبقة.

ب- أن يكون للطبقات ذات التفاوت الأكبر بين عناصرها حجم عيّنة أكبر.

ج- أن يكون للعيّنات ذات التكلفة الأكبر حجم أصغر من تلك العيّنات ذات التكلفة الأقل.

لتوضيح ذلك لنفترض أنّنا نريد معرفة متوسط الدخل السنوي في بلد ما X وذلك بطريقة العيّنة الطبقية، فعندئذ نقوم بتجزئة المجتمع إلى طبقات، ولتكن على سبيل المثال

أ- طبقة الأشخاص الذين لهم متوسط دخل سنوي أقل من 1000 وحْدَة نقدية.

ب- طبقة الأشخاص الذين لهم متوسط دخل سنوي ما بين 1000 و1500 وحْدَة نقدية.

جـ- طبقة الأشخاص الذين لهم متوسط دخل سنوي ما بين 1500 و2000 وحْدَة نقدية.

د - طبقة الأشخاص الذين لهم متوسط دخل سنوي أكثر من 2000 وحْدَة نقَدية.

فلو افترضنا أنَّ هذه الطبقات على الترتيب هي $S_1$، $S_2$، $S_3$ و $S_4$، فعندئذ نقوم بسحب عيِّنات بسيطة من كل طبقة من هذه الطبقات لحساب متوسط الدخل فيها، فإذا لاحظنا أنَّه يوجد تفاوت واضح في حجم الطبقات السابقة، فعندئذ يجب أنَّ نأخذ ذلك بالحسبان لدى تحديد حجم العيِّنة المسحوبة من كل طبقة بحيث يكون متناسباً طرداً مع حجم تلك الطبقة، ويفضَّل استخدم ثابت التناسب المذكور آنفًا نفسه من أجل كل العيّنات، فإذا وجدنا (على سبيل المثال) أنَّ حجم هذه الطبقات هو:

$$N_1 = 3\times10^6 \ \ \& \ \ N_2 = 4\times10^6 \ \ \& \ \ N_3 = 2\times10^6 \ \ \& \ \ N_4 = 10^6$$

فعندئذ يُفضَّل أخذ حجم العيّنات كنسبة مئوية من تعداد كل طبقة، كأنَّ نأخذ 0.01% من كل طبقة، فيكون حجم العيّنة الواجب سحبها من الطبقة $S_1$، $S_2$، $S_3$ و $S_4$ على الترتيب هو $n_1 = 300$، $n_2 = 400$، $n_3 = 200$ و $n_4 = 100$، ومن ثمَّ يُحسب متوسط الدخل السنوي للفرد في كل عيّنة من هذه العيّنات، ثمَّ أخذ المتوسط للقيم التي حصلنا عليها، فتكون هذه القيمة الأخيرة هي (على وجه التقريب) الممثِّل لمتوسط الدخل السنوي للفرد في هذا المجتمع وفقاً للمعاينة الطبقية.

### (٤ , ١ , ٢ , ٩) العيّنة العنقودية Cluster Sample

يقوم الباحث بتكوين العيّنة في هذا النوع من العيّنات على عدّة مراحل، وذلك على النحو الآتي:

١- تقسيم المجتمع إلى فئات متجانسة قدر الإمكان، وهذا التقسيم إمَّا أنَّ يكون افتراضياً أو أنَّ يكون مفروضاً وفقاً لطبيعة المسـألة ونوعية المجتمع المدروس.

٢- تقسيم الفئات السابقة إلى فئات جزئية متجانسة قدر الإمكان، وهذا التقسيم إمَّا أنَّ يكون افتراضياً أو أنَّ يكون مفروضاً وفقاً لطبيعة المسـألة، ومن الممكن أن يكون هناك تقسيمات جزئية أخرى لهذه الفئات الجزئية الأخيرة أيضاً.

٣- أخذ عيّنة عشوائية (بسيطة أو منتظمة) من كل فئة من هذه الفئات الجزئية الأخيرة.

٤- دمج العيّنات الأخيرة في عيّنة واحدة لتشكل العيّنة المطلوبة.

لتوضيح ذلك لنفترض أنّنا نريد دراسة الوضع المعاشي في مدينة ما X مع ضواحيها بطريقة المعاينة العنقودية.

نقوم أولاً بتقسيم المجتمع إلى فئات متجانسة قدر الإمكان، كأن نضع السكان المقيمين في مركز المدينة في فئة (طبقة) $C_1$، والسكان المقيمين على أطراف المدينة في فئة (طبقة) $C_2$، وسكان الأرياف القريبة ومتساوية البعد عن المدينة في فئة (طبقة) $C_3$، وهكذا مبتعدين عن المدينة حتى سكان الأرياف الأكثر بعداً ومتساوية البعد عن المدينة في فئة (طبقة) أخيرة $C_k$. بعد ذلك نقوم بتجزئة كل فئة عدد من الفئات (على سبيل المثال إلى ثلاث فئات: الموظفين، والمقاولين، وأصحاب المهن الحرّة)، ومن الممكن أن نلجأ إلى تقسم كل فئة من الفئات الأخيرة إلى فئات جزئية أخرى (كأن نقسِّم فئة الموظفين إلى موظفين الدرجة الأولى، وموظفين الدرجة الثانية، ... وموظفين الدرجة الخامسة)، وبالمثل بالنسبة إلى بقية الفئات إن كان ذلك ممكناً وواجباً من أجل الدراسة. بعد ذلك نسحب عيّنة عشوائية (بسيطة أو منتظمة) من كل فئة من الفئات الأخيرة، وأخيراً نقوم بدمج العناصر التي حصلنا عليها في فئة واحدة لتكوِّن العيّنة العشوائية العنقودية المطلوبة.

بالطبع من الممكن أن تتمّ عملية المعاينة هذه على عدد أكبر من المراحل كأن يكون هناك تجزئة لكل من فئات الموظفين (كأن نجزئها إلى: موظفين ميدانيين، وإداريين، و...) والمقاولين (كأن نجزئها إلى: مقاولين للبناء، ومقاولين للطرق، ومقاولين للكهرباء، و...) وأصحاب المهن الحرَّة (كأن نجزئها إلى: مهنة السباكة، ومهنة إصلاح السيارات، ومهنة النجارة و...) أيضاً.

نكتفي بهذا القدر من التقديم حول بعض النماذج من العيّنات العشوائية؛ وذلك بسبب كثرة نماذجها التي لا يمكن تغطيتها في هذه العجالة من التقديم حيث يمكن لمن يود الاطلاع على المزيد الرجوع إلى المراجع ذات الصلة.

### (٩،٢،٢) العيّنات العمديّة (أو القصديّة) Intentional Samples

إنَّ العيّنات العمديّة هي عيِّنات غير عشوائية (وكما هو واضح من تسميتها) تقوم على التدخل الشخصي للباحث في انتقاء عناصر العيّنة معتمداً في ذلك خبراته وحدسه، ولهذا لا يمكننا استخدام نظرية الاحتمالات والإحصاء الرياضياتي في تعميم النتائج التي نتوصل إليها من العيّنة على المجتمع، ولكن هذا لا يعني أنَّها عديمة الفائدة وذلك لأنَّه من الممكن أنَّ تلعب دوراً مهّاً في بعض الحالات. وفيما يلي نقدِّم بعض هذه النماذج، ولكن بشكل موجز وسريع، ومن أهمها:

#### (٩،٢،٢،١) عيِّنات الحصص Quota Samples:

في هذا النوع من العيّنات يقوم الباحث بانتقاء فئات معيّنة من المجتمع، وبحيث يُحدَّد لكل فئة حصة معيّنة لأجل اختيار عناصر العيّنة التي سيتم دراستها، وهكذا نلاحظ أنَّ هذا النوع من العيّنات يشابه نموذج العيّنات العشوائية الطبقية باستثناء غياب عامل العشوائية، فعلى سبيل المثال لدى التحقيق في قضية ما يوعَز لكل مُحقِّق جمع معلومات من عدد محدَّد سلفًا من الأفراد، ولكن يتّم ترك اختيار الأفراد وفقاً لطلب المُحقِّقين، فعندئذ العيّنة التي خضعت للتحقيق هي من نوع عينات ًالحصص.

#### (٩،٢،٢،٢) العيّنات المنتقاة Selected Samples

في هذا النموذج من العيّنات يقوم الباحث بانتقاء عناصر معيّنة من المجتمع اعتقاداً منه أنَّ هذه العناصر مُمثِّلة للمجتمع بشكل جيد، وهذا الانتقاء للعناصر ليس له مقياس مُحدَّد إذ إنَّ ذلك يعتمد على خبرة ومهارة الباحث. هذا وتجدر الملاحظة هنا إلى أنَّ هذا النوع من العيّنات قد يعطي نتائج أكثر دقة من النموذج السابق عندما يتعلق الأمر بتقدير المتوسط للظاهرة المدروسة، في حين أنَّ النتائج المتعلّقة بتشتت الظاهرة التي قيد الدرس ستكون أقل إرضاءً مما هو عليه الحال لدى دراسة المتوسط.

#### (٩،٢،٢،٣) العيّنات الكبيرة الحجم High-Size Samples

إنَّ هذا النموذج من العيّنات مبني على أساس سحب عيِّنات كبيرة بالقدر المستطاع، وذلك بغض النظر عن نوعية العناصر المسحوبة من المجتمع، ويبرر الباحثون هذا الأسلوب من خلال الكم الكبير للمعلومات التي سيحصلون عليها من العيّنات التي يُفترَض أنَّ تعطي تصوراً أكثر دقَّة بخصوص الظاهرة المدروسة. بالطبع هذا الطرح يكون مقبولاً إلى حدٍّ ما عندما يكون المجتمع محدوداً وحجم العيّنة قريب من حجم المجتمع.

توجد أنواع أخرى كثيرة من العيّنات غير العشوائية مثل عينات **كرة الثلج** و **العينات الملائمة** و... حيث يمكن لمن يود الاطلاع على المزيد الرجوع إلى المراجع التي تبحث في نظرية المعاينة.

#### (٩،٢،٢،٣،١) ملاحظات

١- يجب الحذر لدى التعامل مع نتائج العيّنات غير العشوائية، وذلك بسبب عامل التحيُّز الذي يمكن أنَّ يحدث من قبل الباحث.

٢- هنا نودّ الإشارة إلى أنَّه لا يوجد مقياس لتحديد حجم العيّنة المسحوبة إن كان كبيراً أو صغيراً، إذ إنَّ ذلك يتوقف على طبيعة المجتمع المدروس، ونوعية الظاهرة التي قيد الدرس، وكذلك على الإمكانيات المتوفرة لدى الباحث، ولكن اتفقنا مجازياً على أنَّ العيّنات التي يكون حجمها $n \geq 32$ تعدّ من فئة العيّنات الكبيرة (وضّحنا السبب في الفصل الأول)، وأمّا العيّنات التي يكون حجمها $n < 32$ فإنَّها تُصنَّف كعيِّنات صغيرة، ولكنَّ هذا العدد/ 32 / المتَّفق عليه قابل للتعديل أحياناً حسب طبيعة المسألة المدروسة ومتطلباتها، وكذلك وفقاً للإمكانيات المتوفرة لدى الباحث.

٣- سنركز دراستنا المقبلة على استخدام العيّنات العشوائية (والبسيطة منها خصوصاً)، ومن ثَّم يجب أنَّ تتحقَّق الشروط الآتية في كلٍّ من المجتمع والعيِّنة:

أ- يجب أن يكون المجتمع الذي ستسحب منه العيّنة متجانساً إلى أقصى حدٍّ ممكن بخصوص المسألة المدروسة.

ب- يجب أن تكون جميع عناصر المجتمع الذي ستسحب منه العيّنة مستقلّة بعضها عن بعض (لا تأثير لعنصر على آخر).

جـ- يجب أنَّ يتم سحب عناصر العيّنات من المجتمع بطريقة عشوائية (دون التحيّز لعنصر على حساب عنصر آخر).

٤- فيما يلي، وعندما نذكر كلمة "**عيّنة**"، فإنَّنا نقصد بذلك عيّنة عشوائية ما لم نشير إلى خلاف ذلك صراحةً، وكذلك عندما نذكر كلمة مجتمع فإنَّنا نقصد بذلك أنّه **مجتمع إحصائيّ**.

## (٩,٣) البناء الرياضياتي النظري للعينة العشوائية البسيطة

### Random Simple Sample

لقد ذكرنا فيما سبق الشروط الواجب توفرها حتى نحصل على عيّنة بسيطة، ولكن على أرض الواقع يُنظر بعين الشك إلى النتائج التي نحصل عليها من سحب عيّنة واحدة فقط، وذلك لأنّه ما يدرينا أنّنا سنحصل على النتائج نفسها من عيّنة أخرى لها الحجم نفسه وتحت الشروط نفسها أيضاً؟

لتقديم إجابة على السؤال السابق يُلجأ عادة إلى سحب عدد (كبير أو صغير، ذلك حسب الإمكانيات المتوفرة لدى الباحث) من العيّنات العشوائية ذات الحجم المتساوي، ومن ثمَّ أخذ النتيجة المتمخّضة عن جميع هذه العيّنات كنتيجة لعيّنة عشوائية ذات حجم مساوٍ لحجم كل عيّنة من العيّنات السّابقة، ولتوضيح ما سبق سنفترض أنّنا نريد معرفة متوسط الطول لمجتمع طلبة في جامعة ما بوساطة أخذ عينات بسيطة، فعندئذ نفترض أنَّ $\Omega$ هي مجموعة كل الطلبة في هذه الجامعة (أي إنّ $\Omega$ تمثّل مجتمع الطلبة في هذه الجامعة)، وأنَّ لجميع الطلاب النصيب نفسه في الاختيار، وأنّه لا تأثير لطالب على آخر بخصوص الظاهرة التي نودّ دراستها (قياس طول الطالب). عندئذ نقوم بسحب عيّنة $\{\omega_{11}, \omega_{12}, ..., \omega_{1n}\}$ (حجمها يساوي $n$) من مجتمع الطلبة، ومن ثمَّ نقوم بقياس طول كل طالب من هذه العيّنة بوساطة مسطرة (مرقّمة بالسنتيمتر مثلاً) سنرمز لها بـ $X$، فنحصل نتيجة لذلك على مجموعة من القيم العشوائية $\{x_{11}, x_{12}, ..., x_{1n}\}$ لأنّه لا يمكننا معرفة طول الطالب الذي سيخضع للقياس بشكل مُسبق وبدِّقة، علماً أنَّ $x_{1i}$ مع $i \ni \mathbb{N}_n$ هي القيمة الناتجة عن قياس طول الطالب $\omega_{1i}$ (ذي الرقم $i$ في هذه العيّنة). بعد ذلك نسحب من مجتمع الطلبة عيّنة ثانية $\{\omega_{21}, \omega_{22}, ..., \omega_{2n}\}$ (حجمها يساوي $n$ أيضاً، والسحب قد يكون مع الإعادة أو بدون إعادة إذا كان حجم المجتمع كبيراً)، ومن ثمَّ نقوم بقياس طول كل طالب من هذه العيّنة فنحصل نتيجة لذلك على مجموعة من القيم العشوائية $\{x_{21}, x_{22}, ..., x_{2n}\}$ علماً أنَّ $x_{2i}$ مع $i \ni \mathbb{N}_n$ هي نتيجة قياس طول الطالب $\omega_{2i}$ (ذي الرقم $i$ في العيّنة الثانية). ثمَّ نكرّر هذه العملية من سحب العيّنات لـ $m$ مرّة، فنحصل من سحب العيّنة $\{\omega_{m1}, \omega_{m2}, ..., \omega_{mn}\}$ على مجموعة من القيم العشوائية $\{x_{m1}, x_{m2}, ..., x_{mn}\}$ علماً أنَّ $x_{mi}$ مع $i \ni \mathbb{N}_n$ هي نتيجة قياس طول الطالب $\omega_{mi}$ (ذي الرقم $i$ في العيّنة الأخيرة). فلو قمنا الآن بصبّ النتائج السّابقة في جدول كما يلي:

الجدول (٩,١)

| | | Observation الملحوظة | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | $X_1$ | $X_2$ | $X_3$ | $\cdots$ | $X_n$ | | |
| Result of | First Sample | $x_{11}$ | $x_{12}$ | $x_{13}$ | $\cdots$ | $x_{1n}$ | لعيّنة الأولى | نتائج |
| | Second Sample | $x_{21}$ | $x_{22}$ | $x_{23}$ | $\cdots$ | $x_{2n}$ | لعيّنة الثانية | |
| | $\vdots \quad \vdots$ | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots \ \vdots \ \vdots$ | $\vdots$ | $\vdots \ \vdots$ | |
| | Last Sample | $x_{m1}$ | $x_{m2}$ | $x_{m3}$ | $\cdots$ | $x_{mn}$ | لعيّنة الأخيرة | |

ووضعنا $x_i := g(x_{1i}, x_{2i}, ..., x_{mi})$ من أجل كل $i \ni \mathbb{N}_n$ بحيث تُأخذ $g$ كدالّة مناسبة لهدف الدراسة الإحصائية (كأنَّ تكون $g$

دالّة المتوسط مثلاً)، فعندئذ يمكن النظر بعين الرضى إلى العيّنة التي مجموعة نتائجها $\{x_1, x_2, ..., x_n\}$ كتمثيل جيد للمجتمع من أجل الظاهرة المطلوب دراستها، وذلك لأنّها حصيلة العديد من العيّنات العشوائية المسحوبة من المجتمع نفسه ولها الحجم نفسه.

الآن لنَرجع قليلاً إلى الشرح السابق لننظر ما الذي يمكننا استنتاجه؟

في الحقيقة نستنتج ممّا سبق ما يلي:

١- إنَّ المتغيِّر $X$ (المُمثَّل بالمسطرة) هو متغيِّر عشوائي يصف المجتمع $\Omega$ بخصوص الظاهرة التي قيد الدراسة (فعلى سبيل المثال لو كان المجتمع طبيعياً فعندئذ سيكون لـ $X$ توزيع طبيعي).

٢- إنَّ القيم الواقعة في الصف الأول من الجدول السابق هي قيم ناتجة عن إخضاع عناصر العيّنة الأولى لهذا المتغيِّر العشوائي $X$، وهذا النقاش ساري المفعول من أجل بقية القيم الواقعة في الصفوف الأخرى من الجدول السابق.

٣- إنَّ العناصر الموجودة في العمود الأول من ذلك الجدول يمكن النظر إليها على أنّها قيم لمتغيِّر عشوائي $X_1$، وكذلك يمكن النظر إلى قيم العمود الثاني، والثالث، ... والأخير على أنّها قيم لمتغيِّر عشوائي $X_2$، $X_3$، ... و $X_n$ على الترتيب.

٤- المتغيِّرات العشوائية $X_1$، $X_2$، ...و $X_n$ مستقلّة عشوائياً، وهذا الاستقلال ناتج عن استقلال عناصر المجتمع بعضها عن بعض.

٥- المتغيِّرات العشوائية $X_1$، $X_2$، ...و $X_n$ متطابقة في التوزيع مع المتغيِّر العشوائي $X$ الواصف للمجتمع (الذي لعب دور أداة القياس)، وهذا التطابق حاصل من كون قيم هذه المتغيِّرات العشوائية ناتجة عن قيم المتغيِّر العشوائي $X$ نفسه.

هكذا نرى أنَّ القيم الناتجة عن كل عيّنة من العيّنات السابقة هي قيم لمتَّجه عشوائي سنرمز له بـ $\mathfrak{X}_n = (X_1, X_2, ..., X_n)$، ومُركّباته تُحقِّق الخواص /٤/ و/٥/ السابقة، ومن أجل $n$ قيمة مُثبّتة فإنَّ المتّجه العشوائي $\mathfrak{X}_n$ يُطبِّق على المجموعة $\Omega^n$ وفقاً للعلاقة الآتية:

$$\mathfrak{X}_n : \Omega^n \longrightarrow \mathbb{R}^n \;\; ; (\omega_1, \omega_2, ..., \omega_n) \mapsto \mathfrak{X}_n(\omega_1, \omega_2, ..., \omega_n) := (X_1(\omega_1), X_2(\omega_2), ..., X_n(\omega_n))$$

### (١, ٣, ٩) ملاحظات

١-من الممكن أنَّ يكون عدد العيّنات العشوائية المسحوبة من المجتمع غير منته، ومن ثمَّ تكون قيم المتغيِّرات العشوائية $X_1$، $X_2$، ...و $X_n$ غير منتهية أيضاً.

٢- إنَّ المتغيِّرات العشـوائية $X_1$، $X_2$، ...و $X_n$ تُدعى **ملحوظات** (أو مشاهدات)، وأمّا القيم الناتجة عن هذه المتغيِّرات العشوائية فإنَّها تُدعى **القيم الملحوظة** (أو المشاهدة) Observation Values للعيِّنة، ويُرمز لهذه القيم بأحرف صغيرة من قبيل $x_1$، $x_2$، ...و $x_n$ (على غرار ما استخدم في الاحتمالات).

٣- إنَّ العيّنة العشوائية $\mathfrak{X}_n = (X_1, X_2, ..., X_n)$ التي تمَّ بناؤها وفقاً للعرض السابق تُدعى **عيّنة عشوائية بسيطة**.

٤- عندما نذكر مستقبلاً أنّه لدينا عيّنة $\mathfrak{X}_n$ فإنَّنا نقصد بذلك أنّها عيّنة عشوائية بسيطة حجمها يساوي $n$، وأنَّ المتَّجه $(x_1, x_2, ..., x_n)$ الممثِّل لقيمة $\mathfrak{X}_n$ إنّما هو نتيجة لتطبيق $\mathfrak{X}_n$ على متّجه $(\omega_1, \omega_2, ..., \omega_n)$ (الذي يمثِّل عيّنة حجمها $n$ ستطبَّق عليها الدراسة الإحصائية) من $\Omega^n$ وفقاً للعلاقة:

$$\mathfrak{X}_n(\omega_1, \omega_2, ..., \omega_n) := (X_1(\omega_1), X_2(\omega_2), ..., X_n(\omega_n))$$

ولذلك حتى لو لم تظهر عناصر العيّنة $(\omega_1, \omega_2, ..., \omega_n)$ التي ستطبَّق عليها الدراسة في عرض العيّنة العشوائية $\mathfrak{X}_n$ فإنَّه يجب ألا يغيب عن ذاكرتنا أنّها موجودة ضمناً في هذه الصيغة.

٥- في حا ل عدم ذكر حجم المجتمع أو عدم التنويه إلى ما يُدلّل على ذلك فإنّنا نقصد بذلك مجتمعاً غير محدود أو كبير جداً بحيث إن عملية سحب العيّنات لا تؤثّر في حجمه، وذلك بغض النظر عن الطريقة التي يتمّ فيها سحب العيّنة سواء كانت مع الإعادة أم بدون إعادة.

٦- عندما نذكر أنّه لدينا مجتمع موصوف من خلال متغيّر عشوائي $X$ فإنّما نقصد بذلك أنّ التوزيع الاحتمالي لهذا المجتمع هو توزيع $X$ نفسه، وأنّ أي عيّنة $\mathcal{X}_n = (X_1, X_2, ..., X_n)$ مسحوبة من هذا المجتمع ستكون مولّدة وفقاً للطريقة التي ذكرت سابقاً، ومن ثمّ يكون لجميع الملحوظات $X_1$، $X_2$،...و $X_n$ توزيع المتغيّر العشوائي $X$ نفسه أيضاً.

٧- عندما نذكر أنّ للمتغيّر العشوائي $X$ (الواصف للمجتمع) دالّة كثافة احتمالية $f_X\ (x;\theta)$ فإنّنا نقصد بذلك ما يلي:

أ- إذا كان $X$ متغيّراً عشوائياً مستمرًّا فإنّ $f_X\ (x;\theta)$ هي دالّة الكثافة الاحتمالية نفسها التي عُرِّفت سابقاً من أجل المتغيّرات العشوائية المستمرّة، وأنّ القيمة $\theta$ هي مَعْلَمة التوزيع الذي من الممكن أن يكون متّجهاً ذا بُعد منته.

ب- إذا كان $X$ متغيّراً عشوائياً متقطّعاً بمجموعة قيم $\mathbf{X} = \{x_i\ ;\ i \in I\}$ فإنّ $f_X\ (x;\theta)$ هي دالّة الكتلة الاحتمالية المعروفة سابقاً من أجل المتغيّرات العشوائية المتقطّعة، ولكنّها كدالّة بـ $x_i$ و $\theta$ معاً، والمثالان الآتيان يوضّحان هذه الملاحظة.

**(٩، ٣، ٢) أمثلة**

١- ليكن لدينا مجتمعاً طبيعياً بمتوسط $\mu$ وانحراف معيّاري $\sigma$، فعندئذ يكون للمتغيّر العشوائي $X$ الواصف لهذا المجتمع توزيع $\mathbf{N}\ (\mu, \sigma)$، وأمّا دالّة كثافته الاحتمالية $f_X\ (x;\theta)$ فلها العرض الآتي من أجل كل $x \in \mathbb{R}$:

$$f_X\ (x;\theta) = f_X\ (x;(\mu,\sigma)) = \frac{1}{\sigma\sqrt{2\pi}}\ \mathbf{exp}\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

وهذا يعني أنّ المَعْلمة $\theta = (\mu, \sigma)$ هو متّجه من $\mathbb{R}\times(0,\infty)$.

٢- ليكن لدينا مجتمعاً برنولياً مَعْلمته $p$ (احتمال النجاح)، فعندئذ يكون للمتغيّر العشوائي $X$ الواصف لهذا المجتمع توزيع ثنائي النقطة في $x_1 = 1$ و $x_2 = 0$ (توزيع برنولي)، وأمّا دالّة كثافته الاحتمالية $f_X\ (x;\theta)$ فلها العرض:

$$f_X\ (x;\theta) = f_X\ (x;p) = P\ (X = x) = p^x\ (1-p)^{1-x} \quad ; x = 0,1$$

فنلاحظ هنا أنّ المَعْلمة $\theta = p$ هو قيمة من الفترة $(0, 1)$.

## (٩،٤) النموذج الرياضياتي النظري لمجتمع إحصائي

**Theoretical Mathematical Model for Population**

لقد رأينا في فصول سابقة أنّه من أجل كل تجربة عشوائية يمكننا إيجاد نموذج رياضياتي نظري يمثّله ألا وهو الفضاء الاحتمالي $[\Omega, \mathcal{A}, P]$ الموافق لهذه التجربة العشوائية. كذلك الأمر بالنسبة إلى المجتمع، فمن أجل أي مجتمع يمكننا تعيين نموذج رياضياتي نظري يمثّله، فلو كانت $\Omega$ (مجموعة غير خالية) هي مجموعة العناصر لمجتمع ما، و $\mathbb{S}$ هو $\sigma$-**جبر** فوق $\Omega$، وأخيراً نأخذ $\mathcal{P} = \{P_\theta; \theta \in \Theta \subseteq \mathbb{R}^k\}$ أسرة قياسات احتمالية فوق $[\Omega, \mathbb{S}]$، علماً أنّ $k \in \mathbb{N}$ مثبّت و $\theta$ متّجه من $\mathbb{R}^k$ ويُدعى **مَعْلَمة المجتمع**.

تحت هذه الافتراضات يُطلق على الثلاثية $[\Omega, \mathbb{S}, \mathcal{P}]$ اسم **النموذج الرياضياتي النظري للمجتمع الإحصائي**.

**(٩، ٤، ١) ملاحظات**

١- مما سبق نلاحظ أنّ $\theta$ يمثّل مَعْلَمة المجتمع، ويمكن أن يكون متّجهاً من $\mathbb{R}^k$ عندما يكون $2 \le k$.

٢- إنَّ مجموعات عناصر العيّنات التي ستخضع للدراسة هي عناصر من الـ σ- **جبر** $\mathbb{S}$ .

٣- على سبيل التبسيط والاختصار سنستخدم عبارة "**مجتمع** $[\Omega, \mathbb{S}, \mathcal{P}]$" عوضاً عن عبارة مجتمع بنموذج رياضياتي نظري $[\Omega, \mathbb{S}, \mathcal{P}]$ ما لم يؤدي ذلك إلى التباس في المعنى أو المدلول.

٤- نعلم أنَّه في النموذج الرياضياتي النظري $[\Omega, \mathcal{A}, P]$ لتجربة عشوائية يكون للقياس الاحتمالي $P$ صيغة مُحَدَّدة (توزيع معيَّن)، في حين أنَّنا نجد في النموذج الرياضياتي النظري للمجتمع أنَّ القياسات الاحتمالية $P_\theta$ تابعة لقيم المَعْلمة $\theta$ ، ومن ثمَّ يجب جمع كل ما هو ممكن من معلومات حول $\theta$ لكي يصبح $P_\theta$ قياساً احتمالياً مناسباً للمجتمع المُعطى، والأمثلة الآتية توضِّح لنا ذلك.

أ- ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً طبيعياً بمَعْلمتين $\mu$ و $\sigma$ ، أي إنَّ هذا المجتمع موصوف من خلال متغيِّر عشوائي $X$ خاضع للتوزيع الطبيعي $\mathbf{N}(\mu, \sigma)$ ، فعندئذ يكون لدينا:

$$\mathcal{P} = \{P_\theta ; \theta \in \Theta\} = \{P_{(\mu, \sigma)} ; (\mu, \sigma) \in \mathbb{R} \times (0, \infty)\}$$

ومن ثمَّ يجب البحث في معرفة قيمة كل من المَعْلمتين $\mu$ و $\sigma$ لكي يصبح $[\Omega, \mathbb{S}, \mathcal{P}]$ نموذجاً رياضياتيًّا مناسباً للمجتمع المُعطى. إنَّ قيمة كل من $\mu$ و $\sigma$ مجهولة في الحالة العامّة ولذلك علينا تقدير كلّ من $\mu$ و $\sigma$ أو أحدهما إن كان الآخر مُعطى، علماً أنَّ عملية التقدير هذه تتمُّ باستخدام عيِّنات عشوائية تُسحب من المجتمع.

ب- ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً أسيًّا بمَعْلَمة $\lambda > 0$ ، أي إنَّ هذا المجتمع موصوف من خلال متغيِّر عشوائي $X$ خاضع للتوزيع الأسّي بمَعْلَمة $\lambda > 0$ ، فعندئذ يكون:

$$\mathcal{P} = \{P_\theta ; \theta \in \Theta\} = \{P_\lambda ; \lambda \in (0, \infty)\}$$

ومن ثمَّ يجب علينا معرفة قيمة $\lambda$ أو تقديرها من عيِّنات عشوائية تُسحب من المجتمع بحيث يصبح $[\Omega, \mathbb{S}, \mathcal{P}]$ نموذجاً رياضياتياً مناسباً لهذا المجتمع المُعطى.

٥- يتبيَّن لنا مما تقدَّم (في هذا الفصل وفي الفصول السابقة) أنَّ الغاية من أخذ عيّنة من المجتمع المفترض قد يكون لها أكثر من هدف، ومنها على سبيل المثال لا الحصر:

أ- الحصول على معلومات حول توزيع هذا المجتمع الذي يعبَّر عنه بوسائط عدَّة منها على سبيل المثال جدول التوزيع التكراري، المدرَّج التكراري، المضلَّع التكراري أو....

ب- معرفة بعض القيم الكميّة المميِّزة المتعلِّقة بالمجتمع مثل: المتوسط، التباين، الالتواء أو...

إنَّ عملية الحصول على هذه المعلومات تتمّ عادة من خلال تقديرها بقيم سُتحصل عليها من العيّنات العشوائية المسحوبة من المجتمع المفترض، ولكنَّ الوصول إلى هذه الغاية بشكل علمي معقول يتطلب منَّا تقديم بعض التعاريف والدراسات التي سنقدِّم بعضها على مدى القسم المتبقي من هذا الفصل.

## (٩,٥) الإحصاءات ودوال توزيع العينات

### Statistics and distribution functions of samples

يقدِّم لنا التعريف الآتي مفهوم الإحصاءة أو ما يُعرف باسم دالَّة العيّنة Sample Function أيضاً.

**(٩ , ٥ , ١) تعريف (الإحصاءة Statistic)**

ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ ، ولنأخذ $\mathfrak{X}_n$ عيّنة من هذا المجتمع، وبفرض أنَّ $g$ دالَّة حقيقية

معرَّفة على $\mathbb{R}^n$ وقيوسة بالنسبة إلى $\mathfrak{R}$ و $\mathfrak{R}^n$، فعندئذ يُطلق على المتغيِّر العشوائي $T := g(\mathfrak{X}_n)$ اسم **إحصاءة** Statistic أو **دالَّة عيِّنة** Sample Function (سنستخدم في سياق دراستنا المقبلة عبارة "إحصاءة" للدلالة على الدالَّة $g(\mathfrak{X}_n)$).

فيما يلي نقدِّم بعض النماذج الشهيرة من الإحصاءات.

**(١, ١, ٥, ٩) أمثلة**

ليكن $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً، ولنأخذ $\mathfrak{X}_n = (X_1, X_2, ..., X_n)$ عيِّنة مسحوبة من هذا المجتمع، وبفرض أنَّ $g$ دالَّة حقيقية معرَّفة على $\mathbb{R}^n$ وقيوسة بالنسبة إلى $\mathfrak{R}$ و $\mathfrak{R}^n$ (جميع الدوال التي سنقدِّمها في هذا المثال قيوسة بالنسبة إلى $\mathfrak{R}$ و $\mathfrak{R}^n$)، فعندئذ:

١- إذا أخذنا $g$ دالَّة بالعينة $\mathfrak{X}_n$ ومعرَّفة من خلال العلاقة الآتية:

$$g(\mathfrak{X}_n) = \frac{1}{n}\sum_{i=1}^{n} X_i$$

فإنَّ هذه الإحصاءة تُدعى **متوسط العيِّنة** Mean of Sample $\mathfrak{X}_n$، وكنّا قد رمزنا لها سابقاً (انظر العلاقة [8–11-a] في الفصل الثامن) بـ $\bar{X}_n$ حيث كانت قيمة $n$ تتحول على كل قيم $\mathbb{N}$، ولكن بسبب أنَّ قيمة $n$ هنا مثبَّتة (من أجل عيِّنة مُحدَّدة الحجم) فإنَّنا سنستخدم الرمز $\bar{X}$ بدلاً من $\bar{X}_n$ ما لم يؤدِّ ذلك إلى التباس، أي إنَّه لدينا:

$$\bar{X} = \frac{X_1 + X_2 + ... + X_n}{n} \qquad [9,2]$$

وأمّا قيمة هذه الإحصاءة فسوف نرمز لها بـ $\bar{x}$ لكي يتفق مع ما تمَّ تقديمه سابقاً.

٢- إذا أخذنا $g$ كدالَّة بالعينة ومعرَّفة من خلال العلاقة الآتية:

$$g(\mathfrak{X}_n) = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

فعندئذ تُدعى هذه الإحصاءة **تباين العيِّنة** Variance of Sample $\mathfrak{X}_n$ (أو التباين العملي الذي قدِّم في الفصل الثاني من أجل مجموعة من البيانات)، ويُرمز له بـ $S_X^2$، أي إنَّه لدينا:

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 \qquad [9,3]$$

وأمّا قيمة هذه الإحصاءة فيُرمز لها بـ $s_X^2$، وفي حال عدم الوقوع في التباس (كأن تكون الدراسة متعلِّقة بمجتمع وحيد) فإنَّنا سنكتب $S^2$ و $s^2$ عوضاً عن $S_X^2$ و $s_X^2$ وذلك على سبيل التبسيط.

إنَّ الجذر التربيعي الموجب لتباين العيِّنة يُدعى **الانحراف المعياري للعيِّنة** ويُرمز له بـ $S_X$ ولقيمته بـ $s_X$، وفي حال عدم الوقوع في التباس يُكتب $S$ و $s$ فقط على سبيل التبسيط.

٣- إذا أخذنا $g$ دالَّة بالعينة $\mathfrak{X}_n$ ومعرَّفة من خلال العلاقة الآتية:

$$g(\mathfrak{X}_n) = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2$$

وسنرمز لهذه الإحصاءة بـ $\tilde{S}_X^2$، وهنا نلاحظ أنَّ صيغة هذه الإحصاءة توافق الصيغة النظرية للتباين في الاحتمالات، وكذلك توافق صيغة التباين لبيانات مجتمع (كما مرَّ معنا في الفصل الثاني من هذا الكتاب)، وهي بالفعل صيغة تباين للعينة أيضاً (ولكنَّها ليست التباين العملي)، ولكي لا تلتبس هذه التسمية مع تسمية تباين العيِّنة $S^2$ سنكتفي بتسميتها بالإحصاءة $\tilde{S}^2$ فقط، وفي حال عدم الوقوع في التباس فإنَّنا سنكتب $\tilde{S}^2$ فقط، وأمّا قيمة هذه الإحصاءة فسوف نرمز لها بـ $\tilde{s}^2$. أي إنَّه لدينا:

$$\tilde{S}_X^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \bar{X}\right)^2 \qquad [9,4]$$

نشير هنا إلى أنَّ الجذر التربيعي الموجب لـ $\tilde{S}^2$ يُدعى **الانحراف المعياري** أيضاً وسنرمز له بـ $\tilde{S}$، وأمّا قيمته فسوف نرمز لها بـ $\tilde{s}$، وكذلك نلاحظ أنَّ العلاقة التي تربط بين $S^2$ و $\tilde{S}^2$ هي:

$$\tilde{S}^2 = \frac{n-1}{n}S^2 \qquad [9,5]$$

وذلك لأنَّ:

$$\tilde{S}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2 = \frac{n-1}{n}\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2\right] = \frac{n-1}{n}S^2$$

ومن ثَمَّ يكون لدينا:

$$S^2 = \frac{n}{n-1}\tilde{S}^2 \qquad [9,6]$$

٤- إذا أخذنا $g$ دالَّة بالعينة $\mathfrak{X}_n$ ومعرَّفة من خلال العلاقة الآتية:

$$T = g(\mathbb{X}_n) = \frac{1}{n}\sum_{i=1}^{n}X_i^k$$

علماً أنَّ $k \in \mathbb{N}$، فعندئذ هذه الإحصاءة تُدعى **العزم من المرتبة $k$ للعيِّنة** $\mathfrak{X}_n$، وسنرمز لها بـ $\mathbf{M}_X^{(k)}$ وفي حال عدم الوقوع في التباس سنكتب $\mathbf{M}^{(k)}$ فقط على سبيل التبسيط، أي إنَّه لدينا:

$$\mathbf{M}^{(k)} = \frac{X_1^k + X_2^k + \ldots + X_n^k}{n} \qquad [9,7]$$

وأمّا قيمة هذه الإحصاءة فسنرمز لها بـ $\mathbf{m}^{(k)}$، ومن أجل $k = 1$ سنكتب $\mathbf{M}^{(1)} = \mathbf{M}$ وكذلك $\mathbf{m}^{(1)} = \mathbf{m}$.

### (٩,٥,٢) تعريف (دالَّة التوزيع العمليَّة للعيِّنة) Empirical Distribution Function of Sample

ليكن $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، ولنعرِّف من أجل هذه العيِّنة دالَّة حقيقية $F_{n,\omega}$ على $\mathbb{R}$ كما يلي:

$$F_{n,\omega} : \mathbb{R} \longrightarrow [0,1] \;\; ; x \mapsto F_{n,\omega}(x) := \frac{1}{n}\sum_{i=1}^{n}\mathbf{I}_{(-\infty,x)}\left(X_i(\omega_i)\right) \qquad [9,8\text{-a}]$$

فعندئذ يُطلق على الدالَّة $F_{n,\omega}$ اسم **دالَّة التوزيع العمليَّة للعيِّنة** $\mathfrak{X}_n$، علماً أنَّ $\omega = (\omega_1, \omega_2, \ldots, \omega_n)$ هي تلك الأفراد من المجتمع التي طبِّق عليها المتَّجه العشوائي $\mathfrak{X}_n$ في الدراسة الإحصائية.

في الحقيقة يمكن عرض الدالَّة $F_{n,\omega}$ بالشكل المكافئ الآتي أيضاً:

$$F_{n,\omega}(x) = \frac{1}{n}\sum_{i=1}^{n}\mathbf{I}_{(-\infty,x)}(x_i) \qquad [9,8\text{-b}]$$

حيث نلاحظ أنَّ قيمة هذه الدالَّة عند نقطة $x \in \mathbb{R}$ تساوي عدد عناصر العيِّنة التي لها قيم أصغر من $x$ مضروبة بـ $n^{-1}$.

إنَّ المبرهنة الآتية التي سنقبلها دون برهان، وترجع إلى الرياضياتي الروسي **غليفينكو** (وتعرف باسم Glivenko–Cantelli

theorem أيضاً)، تبيِّن أنَّ متتالية دوال توزيع العيِّنة تتقارب من دالَّة توزيع المجتمع (بانتظام) عندما يسعى حجم العيِّنة إلى اللانهاية.

**(٩،٥،٢،١) مبرهنة (غليفينكو)**

ليكن $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ وبدالَّة توزيع احتمالية $F_X$، ولنأخذ $\mathfrak{X}_n$ عيِّنة مسحوبة من هذا المجتمع بدالَّة توزيع عمليَّة $F_{n,\omega}$، فعندئذ تكون العلاقة الآتية محقَّقةً:

$$P\left(\lim_{n\to\infty}\sup_{x\in\mathbb{R}}\left|F_{n,\omega}(x)-F_X(x)\right|\right)=1 \qquad [9,9]$$

**(٩،٥،٢،٢) ملاحظات**

١- من التعريف السابق نلاحظ أنَّه من أجل $x \in \mathbb{R}$ مثبَّتة فإنَّ الدالَّة $F_{n,\bullet}$ هي متغيِّر عشوائي فوق $\Omega$.

٢- من المبرهنة السابقة يتضِّح لنا أنَّ متتالية دوال التوزيع العمليَّة $\left(F_{n,\omega}\right)_{n\in\mathbb{N}}$ متقاربة بانتظام من دالَّة توزيع المجتمع $F_X$ وذلك باحتمال يساوي الواحد.

## (٩،٦) المميزات العددية لبعض الإحصاءات

## Nomerical Characteristics of Some Statistic's

بسبب الأهمِّيَّة الكبيرة لبعض الإحصاءات، وعلى وجه الخصوص الإحصاءات $\bar{X}$، $S^2$ و $\tilde{S}^2$، فإنَّنا سنقوم بتقديم بعض المميِّزات العددية لها من خلال المبرهنة الآتية.

**(٩،٦،١) مبرهنة**

ليكن $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولنفترض أنَّ متوسطه $\mu$ (أي إنَّ $\mathbf{E}X=\mu$) وتباينه $\sigma^2$ (أي إنَّ $\mathbf{var}\,X=\sigma^2$)، ولتكن $\mathfrak{X}_n$ عيِّنة مسحوبة من هذا المجتمع، فعندئذ يكون لدينا ما يلي:

**a)** $\mathbf{E}\,\bar{X}=\mu$ $\qquad$ [9,10]

وهذا يعني أنَّ التوقع الرياضياتي (القيمة الوسطى) للإحصاءة $\bar{X}$ مستقلّ عن حجم العيِّنة.

**b)** $\mathbf{var}\,\bar{X}=\dfrac{\sigma^2}{n}$ $\qquad$ [9,11]

وهذا يعني أنَّ قيمة تباين الإحصاءة $\bar{X}$ تتناسب عكسياً مع حجم العيِّنة $n$، ومن ثَمَّ قيمة تباينه ستتناقص بازدياد حجم العيِّنة $n$.

**c)** $\mathbf{E}\tilde{S}^2=\dfrac{n-1}{n}\sigma^2$ $\qquad$ [9,12-a]

والتي نجد منها أنَّ:

**d)** $\mathbf{E}S^2=\sigma^2$ $\qquad$ [9,12-b]

وهذا يعني أنَّ التوقع الرياضياتي للإحصاءة $\tilde{S}^2$ مرتبط بحجم العيِّنة $n$، في حين أنَّ التوقع الرياضياتي للإحصاءة $S^2$ مستقلّ عن حجم العيِّنة.

**e)** $\mathbf{var}\,\tilde{S}^2=\dfrac{\mu^4-\sigma^4}{n}-2\dfrac{\mu^4-2\sigma^4}{n^2}+\dfrac{\mu^4-3\sigma^4}{n^3}$ $\qquad$ [9,13-a]

وهذا يعني أنَّ قيمة التباين للإحصاءة $\tilde{S}^2$ مرتبطة عكسياً مع حجم العيّنة $n$، ومن ثمَّ قيمة تباينه ستتناقص بازدياد حجم العيّنة $n$ أيضاً.

$$\text{f)}\quad \mathbf{var}\, S^2 = \left(\frac{n}{n-1}\right)^2 \cdot \left(\frac{\mu^4 - \sigma^4}{n} - 2\frac{\mu^4 - 2\sigma^4}{n^2} + \frac{\mu^4 - 3\sigma^4}{n^3}\right) \qquad [9,13\text{-b}]$$

**البرهان:** سنقبل العلاقة [9,13-a] دون برهان، وأمّا العلاقتين [9,12-a] و [9,13-b] فإنَّهما تنتجان عن العلاقتين [9,12-a] و[9,13-b] على الترتيب لدى استخدام العلاقة [9,6] وخواص التباين. أمّا من أجل بقية العلاقات فإنَّه بسبب الاستقلال بين ملحوظات العيّنة $X_1$، $X_2$،...و $X_n$ وتطابقها في التوزيع مع المتغيِّر العشوائي $X$، وكذلك باستخدام خصائص التوقع الرياضياتي والتباين يكون لدينا ما يلي:

$$\textbf{For a)}\quad \mathbf{E}\bar{X} = \mathbf{E}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbf{E}X_i = \frac{1}{n}\, n \cdot \mathbf{E}X = \mu$$

$$\textbf{For b)}\quad \mathbf{var}\,\bar{X} = \mathbf{var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} \mathbf{var}\,X_i = \frac{1}{n^2}\, n \cdot \mathbf{var}\,X = \frac{\sigma^2}{n}$$

$$\begin{aligned}
\textbf{For c)}\quad \mathbf{E}\,\tilde{S}^2 &= \mathbf{E}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right) = \frac{1}{n}\,\mathbf{E}\left[\sum_{i=1}^{n}\left[(X_i - \mathbf{E}X) + (\mathbf{E}X - \bar{X})\right]^2\right] \\
&= \frac{1}{n}\,\mathbf{E}\left[\sum_{i=1}^{n}\left[(X_i - \mathbf{E}X)^2 - 2(X_i - \mathbf{E}X)(\bar{X} - \mathbf{E}X) + (\bar{X} - \mathbf{E}X)^2\right]\right] \\
&= \frac{1}{n}\,\mathbf{E}\left[\sum_{i=1}^{n}(X_i - \mathbf{E}X)^2 - 2(\bar{X} - \mathbf{E}X)\sum_{i=1}^{n}(X_i - \mathbf{E}X) + n(\bar{X} - \mathbf{E}X)^2\right] \\
&= \frac{1}{n}\,\mathbf{E}\left[\sum_{i=1}^{n}(X_i - \mathbf{E}X)^2 - 2(\bar{X} - \mathbf{E}\bar{X})\sum_{i=1}^{n}(X_i - \mathbf{E}X_i) + n(\bar{X} - \mathbf{E}\bar{X})^2\right] \\
&= \frac{1}{n}\,\mathbf{E}\left[\sum_{i=1}^{n}(X_i - \mathbf{E}X)^2 - 2(\bar{X} - \mathbf{E}\bar{X})\left(\sum_{i=1}^{n}X_i - \mathbf{E}\sum_{i=1}^{n}X_i\right) + n(\bar{X} - \mathbf{E}\bar{X})^2\right] \\
&= \frac{1}{n}\left[\sum_{i=1}^{n}\mathbf{E}(X_i - \mathbf{E}X_i)^2 - n\,\mathbf{E}(\bar{X} - \mathbf{E}\bar{X})^2\right] \\
&= \frac{1}{n}\left(n\,\mathbf{var}X_i - n\,\mathbf{var}\bar{X}\right) = \frac{1}{n}\left(n\,\sigma^2 - n\,\frac{\sigma^2}{n}\right) = \frac{n-1}{n}\,\sigma^2
\end{aligned}$$

وبهذا يتمّ البرهان.

### (٩,٦,٢) ملاحظات

١- في الحالة الخاصة، عندما يكون المجتمع المدروس طبيعيًا بمتوسط $\mu$ وانحراف معيّاري $\sigma$، فإنَّه يصبح للعلاقة [9,13-a] العرض الآتي:

$$\mathbf{var}\,\tilde{S}^2 = \frac{2\,(n-1)}{n^2}\sigma^4 \qquad [9,14]$$

٢- تكون الإحصاءتان $\bar{X}$ و $\tilde{S}^2$ (ومن ثمَّ $\bar{X}$ و $S^2$ أيضاً) مستقلتين بعضهما عن بعض إذا وفقط إذا كان المجتمع الذي سُحِبت منه العيّنة العشوائية $\mathfrak{X}_n$ مجتمعاً طبيعياً (انظر [24]).

## (٩,٧) توزيعات المعاينة في حال مجتمع وحيد
## Sampling Distributions for one Population

لقد لاحظنا أنَّ الإحصاءات هي متغيِّرات عشوائية، ومن ثَمّ فمن أهم الأسئلة المتعلِّقة بها هي تلك الأسئلة التي تتعلَّق بمعرفة توزيعات هذه المتغيِّرات العشوائية. إنَّ البحث في التوزيعات الاحتمالية للإحصاءات يُعرف باسم **توزيعات المعاينة**. في الحقيقة يمكننا أن نوجز مفهوم توزيع المعاينة على النحو الآتي:

توزيع المعاينة هو التوزيع الاحتمالي النظري الذي يُبيّن العلاقة الداليّة بين القيم الممكنة لإحصاءة مُعيَّنة استنادًا إلى عيّنة لـ $n$ من الحالات والاحتمال الموافق لكل قيمة، وذلك من أجل جميع العيّنات الممكنة ذات الحجم $n$ والمسحوبة من مجتمع معيَّن، وفي هذا الصدد سنناقش حالتين من هذا الموضوع، وهما:

١- إذا كانت الدراسة تتعلَّق بمجتمع وحيد.

٢- إذا كانت الدراسة تتعلَّق بمجتمعين مستقلَّين.

وعلاوة على ذلك ستكون دراستنا المقبلة مركَّزة على المجتمعات الطبيعية والبرنولية. علماً أنَّنا لن نخوض في تفصيلات الاستنتاجات لهذه التوزيعات، حيث سنكتفي بتقديمها، ولمن يودّ الاطلاع على تفصيلاتها يمكنه الرجوع إلى بعض المراجع التي ذكرناها في بداية هذا الفصل.

### (٩,٧,١) توزيعات المعاينة لمتوسط وتباين عينة من مجتمع طبيعي:

سنفترض على مدى هذه الفقرة أنّه لدينا مجتمعاً $[\Omega, \mathbb{S}, \mathcal{P}]$ طبيعياً (أو طبيعياً على وجه التقريب) بمتوسط $\mu$ وانحراف معيّاري $\sigma$، أي إنَّ هذا المجتمع موصوف من خلال متغيِّر عشوائي $X \sim \mathbf{N}(\mu,\sigma)$، حيث نعلم أنَّ هذين الوسيطين $\mu$ و $\sigma$ هما مَعْلمتي التوزيع الطبيعي، ولذلك عند دراسة توزيعات المعاينة لهذا المجتمع سوف ينصبّ الاهتمام على توزيعات المعاينة لمتوسط العيّنة $\bar{X}$ وتباينها $S^2$ (أو انحرافها المعياري $S$).

#### (٩,٧,١,١) توزيعات المعاينة لمتوسط العيّنة $\bar{X}$:

ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً طبيعياً (أو طبيعياً على وجه التقريب) بمتوسط $\mu$ وانحراف معيّاري $\sigma$، ولنأخذ $\mathfrak{X}_n = (X_1, X_2, ..., X_n)$ مع $n \in \mathbb{N}$ عيّنة مسحوبة من هذا المجتمع. عندئذ من أجل دراسة توزيع المعاينة لمتوسط العيّنة $\mathfrak{X}_n$ سنناقش الحالتين الآتيتين:

١- إذا كان تباين المجتمع معلوماً ويساوي $\sigma_o^2$، أي إنَّ $\sigma^2 = \sigma_o^2$، فعندئذ:

أ- إذا كان حجم المجتمع غير محدود، وكان حجم العيّنة $n$ كبيراً بقدر كاف، فإنَّه سيكون لمتوسط العيّنة $\bar{X}$ توزيع طبيعي (أو طبيعي على وجه التقريب وسنفترض أنَّه طبيعي أيضاً)، ومن ثَمَّ بحسب المميِّزات العدَدية للمتغيِّر العشوائي الخاضع للتوزيع الطبيعي $\mathbf{N}(\mu,\sigma)$ ينتج لدينا:

$$\mu_{\bar{X}} = \mu \qquad [9,15]$$

$$\sigma_{\bar{X}} = \frac{\sigma_o}{\sqrt{n}} \qquad [9,16\text{-a}]$$

ب- إذا كان حجم المجتمع محدوداً ويساوي $N$، وكان حجم العيّنة كبيرًا نسبياً (على الأقل $n > 0.05N$)، فإنَّه سيكون لمتوسط العيّنة $\bar{X}$ توزيع طبيعي على وجه التقريب مع $\mu_{\bar{X}} = \mu$ (كما في العلاقة [9,15]) ولكن بانحراف معيّاري يساوي:

$$\sigma_{\bar{X}} = \frac{\sigma_o}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \qquad [9,16\text{-b}]$$

فلاحظ هنا أنَّه من أجل حجم عيّنة $n$ مثبَّت فإنَّ:

$$\lim_{N\to\infty} \frac{\sigma_o}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{\sigma_o}{\sqrt{n}} \lim_{N\to\infty} \sqrt{\frac{N-n}{N-1}} = \frac{\sigma_o}{\sqrt{n}}$$

ومن ثمَّ تتوافق هذه الصيغة مع الحالة التي يكون فيها حجم المجتمع غير محدود.

٢- إذا كان تباين المجتمع الطبيعي مجهولاً (أي إنَّ ? = $\sigma^2$)، فإنَّه يُبرهن أنَّ للمتغيِّر العشوائي $T = \dfrac{\bar{X} - \mu}{S/\sqrt{n}}$ توزيع ستودنت بـ $n-1$ درجة حُرِّيَّة، ومن ثمَّ بحسب المميِّزات العددية للمتغيِّر العشوائي الخاضع لتوزيع ستودنت ينتج لدينا أنَّه:

أ- إذا كان حجم المجتمع غير محدود، فإنَّه سيكون لـ $\bar{X}$ متوسط $\mu_{\bar{X}} = \mu$ (كما في العلاقة [9,15])، ولكن بانحراف معيّاري يساوي:

$$\sigma_{\bar{X}} = \frac{s}{\sqrt{n}} \qquad [9,17\text{-a}]$$

ب- إذا كان حجم المجتمع محدوداً ويساوي $N$ (على أن يؤخذ $n < 0.05N$ على الأقل)، فإنَّه سيكون ل $\bar{X}$ متوسط $\mu_{\bar{X}} = \mu$ (كما في العلاقة [9,15])، ولكن بانحراف معيّاري يساوي:

$$\sigma_{\bar{X}} = \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} \qquad [9,17\text{-b}]$$

وهنا يُلاحظ أنَّه من أجل حجم عيّنة $n$ مثبَّت سيكون لدينا:

$$\lim_{N\to\infty} \frac{s}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{s}{\sqrt{n}}$$

وهذه الصيغة تتوافق مع الحالة التي يكون فيها حجم المجتمع غير محدود أيضاً.

**(٩,٧,١,٢) ملاحظة**

نودّ أن نشير هنا إلى أنَّ النتائج السابقة بخصوص توزيعات المعاينة لـ $\bar{X}$ في حال مجتمع طبيعي تبقى صحيحة من أجل أي مجتمع بمتوسط $\mu$ وانحراف معيّاري $\sigma$ أيضاً إذا كان حجم العيّنة كبيراً $n \geq 32$، وتتحسن النتائج كلما تزايد حجم العينية أكثر فأكثر. أمّا إذا كان حجم العيّنة صغيراً $n < 32$ فعندئذ تصبح تلك النتائج غير مرضية إلاّ من أجل المجتمعات الطبيعية حيث تحافظ على جودتها.

**(٩,٧,١,٣) توزيعات المعاينة لـتباين العيّنة $S^2$**

ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً طبيعياً بمتوسط $\mu$ وانحراف معيّاري $\sigma$، ولنأخذ $\mathcal{X}_n$ عيّنة مسحوبة من هذا المجتمع، فعندئذ يُبرهن على أنَّه يكون للمتغيِّر العشوائي $\chi = \dfrac{(n-1)S^2}{\sigma^2}$ توزيع كاي مربّع بـ $n-1$ درجة حُرِّيَّة، ومن ثمَّ بحسب المميِّزات العددية للمتغيِّر العشوائي الخاضع لتوزيع كاي مربّع ينتج أنَّ لـ $S^2$ متوسط وتباين كما في العلاقتين الآتيتين على الترتيب:

$$\mu_{S^2} = \sigma^2 \qquad [9,18]$$

$$\sigma^2_{S^2} = \frac{2}{n-1}\sigma^4 \qquad [9,19]$$

**(٩,٧,١,٤) ملاحظة**

نشير هنا إلى أنَّه من أجل مجتمعات كيفية بمتوسط $\mu$ وانحراف معيّاري $\sigma$ يكون للإحصاء $S^2$ توزيعاً احتمالياً غير بسيط (عدا حالة المجتمعات الطبيعية)، حيث يكون لمتوسط وتباين $S^2$ القيمة المعطاة بالعلاقة [9,12-b] و[9,13-b] على الترتيب (انظر في المبرهنة (٩,٦,١) والملاحظة التالية لها (٩,٦,٢)).

**مثال (٩,٧,١,٥)**

لدينا آلة لإنتاج الصفائح الحديدية، فإذا علمت أنَّ أوزان هذه الصفائح تتوزع طبيعياً على وجه التقريب بمتوسط $\mu = 500$ كغ وانحراف معيّاري $\sigma$ ، وأنّنا سحبنا عيّنة $\mathfrak{X}_{169}$ من الإنتاج الكلي لهذه الآلة، فعندئذ:

١- ما هو احتمال ألّا تبتعد قيمة المتوسط لهذه العيّنة عن متوسط مجتمع الصفائح $\mu$ بأكثر من نصف كيلو غرام في كل من الحالتين الآتيتين:

أ- إذا كان $\sigma = 5$ ، ولم يذكر شيء عن حجم الإنتاج الكلي للآلة؟

ب- إذا كان $\sigma = 5$ ، وحجم الإنتاج الكلي للآلة يساوي 1000 صفيحة؟

٢- إذا افترضنا أنّنا لا نعلم شيء عن قيمة الانحراف المعياري لأوزان الصفائح المنتجة (أي إنَّ $\sigma = ?$ )، ووجدنا من العيّنة المسحوبة أنَّ $S = 13$ كغ، فما هو احتمال ألّا تزيد قيمة المتوسط لهذه العيّنة عن متوسط مجتمع الصفائح $\mu$ بأكثر من كيلو غرام واحد؟

٣- إذا افترضنا أنَّ حجم العيّنة $n = 101$ ولدينا $\sigma = 3$ كغ، فما هو احتمال ألّا تزيد قيمة التباين لهذه العيّنة عن قيمة التباين لمجتمع الصفائح $\sigma^2$ بأكثر من 3 كغ؟

**الحل**: من أجل الطلب:

أ-١- بما أنّه لم يذكر شيء عن حجم الإنتاج الكلي للآلة فإنّنا سنفترض أنَّ حجم المجتمع غير محدود، وبملاحظة أنَّ حجم العيّنة كبير بقدر كاف فإنّه سيكون لـ $\bar{X}$ توزيع طبيعي على وجه التقريب بمتوسط $\mu_{\bar{X}} = \mu = 500$ ، وانحراف معيّاري:

$$\sigma_{\bar{X}} = \frac{\sigma_o}{\sqrt{n}} = \frac{5}{13} = 0.385$$

ومن ثمَّ يكون الاحتمال المطلوب هو:

$$P\left(499.5 \le \bar{X} \le 500.5\right) = P\left(\frac{499.5-\mu_{\bar{X}}}{\sigma_{\bar{X}}} \le \frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}} \le \frac{500.5-\mu_{\bar{X}}}{\sigma_{\bar{X}}}\right) = P\left(\frac{499.5-500}{0.385} \le \frac{\bar{X}-500}{0.385} \le \frac{500.5-500}{0.385}\right)$$

والآن بوضع $Z = \dfrac{\bar{X}-500}{0.385}$ واستخدام جدول التوزيع الطبيعي نجد أنَّ:

$$P\left(499.5 \le \bar{X} \le 500.5\right) = P\left(-1.30 \le Z \le 1.30\right) = P\left(Z \le 1.30\right) - P\left(Z < -1.30\right)$$
$$\approx \Phi(1.30) - \Phi(-1.30) = 0.9032 - 0.0968 = 0.8064$$

ب-١- بما أنَّ حجم الإنتاج الكلي للآلة يساوي 1000، وحجم العيّنة كبير بقدر كاف فإنّه سيكون لـ $\bar{X}$ توزيع طبيعي على وجه التقريب بمتوسط $\mu_{\bar{X}} = \mu = 500$ ، وانحراف معيّاري:

$$\sigma_{\bar{X}} = \frac{\sigma_o}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}} = \frac{5}{13}\sqrt{\frac{1000-169}{999}} = 0.351$$

ومن ثمَّ سيكون الاحتمال المطلوب هو:

$$P\left(499.5 \le \bar{X} \le 500.5\right) = P\left(\frac{499.5-\mu_{\bar{X}}}{\sigma_{\bar{X}}} \le \frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}} \le \frac{500.5-\mu_{\bar{X}}}{\sigma_{\bar{X}}}\right)$$
$$= P\left(\frac{499.5-500}{0.351} \le \frac{\bar{X}-500}{0.351} \le \frac{500.5-500}{0.351}\right)$$

والآن بوضع $Z = \dfrac{\bar{X}-500}{0.351}$ واستخدام جدول التوزيع الطبيعي نجد أنَّ:

$$P\left(499.5\le \bar{X}\le 500.5\right)=P\left(-1.42\le Z\le 1.42\right)=P\left(Z\le 1.42\right)-P\left(Z<-1.42\right)$$
$$\approx \Phi(1.42)-\Phi(-1.42)=0.9222-0.0778=0.8444$$

٢- بما أنّنا لا نعلم شيء عن قيمة الانحراف المعياري لأوزان الصفائح المنتجة، وأنّنا وجدنا من العيّنة المسحوبة أنَّ $S=13$، وكذلك بسبب أنّه لم يذكر شيء عن حجم الإنتاج الكلي للآلة فإنّنا سنفترض أنَّ حجم المجتمع غير محدود، وبملاحظة أنَّ حجم العيّنة كبير بقدر كاف فإنّه سيكون لـ $\bar{X}$ توزيع طبيعي على وجه التقريب بمتوسط $\mu_{\bar{X}}=\mu=500$، وانحراف معيّاري:

$$\sigma_{\bar{X}}=\frac{S}{\sqrt{n}}=\frac{13}{\sqrt{169}}=1$$

ومن ثمَّ يكون الاحتمال المطلوب هو:

$$P\left(500\le \bar{X}\le 501\right)=P\left(\frac{500-\mu_{\bar{X}}}{\sigma_{\bar{X}}}\le\frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}}\le\frac{501-\mu_{\bar{X}}}{\sigma_{\bar{X}}}\right)=P\left(500-500\le \bar{X}-500\le 501-500\right)$$

والآن بوضع $Z=\dfrac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}}$ واستخدام جدول التوزيع الطبيعي نجد أنَّ:

$$P\left(500\le \bar{X}\le 501\right)=P\left(0\le Z\le 1\right)=P\left(Z\le 1\right)-P\left(Z<0\right)\approx \Phi(1)-\Phi(0)=0.8413-0.5=0.3413$$

وبهذا يتمّ حل الطلب (٢).

٣- إنَّ الاحتمال المطلوب هو $P\left(9\le S^2\le 12\right)$ حيث لدينا:

$$P\left(\sigma^2\le S^2\le \sigma^2+3\right)=P\left(\frac{(n-1)\sigma^2}{\sigma^2}\le\frac{(n-1)S^2}{\sigma^2}\le\frac{(n-1)(\sigma^2+3)}{\sigma^2}\right)$$

ولأجل حسابه يمكننا أن نكتب الآتي:

$$P\left(9\le S^2\le 12\right)=P\left(\frac{(101-1)9}{9}\le\frac{(101-1)S^2}{9}\le\frac{(101-1)12}{9}\right)$$
$$=P\left(\frac{(101-1)S^2}{9}\le\frac{(101-1)12}{9}\right)-P\left(\frac{(101-1)S^2}{9}<\frac{(101-1)9}{9}\right)$$

وباستخدام جدول توزيع كاي مربّع عند درجة الحرية 100 ينتج لدينا أنَّ:

$$P\left(9\le S^2\le 12\right)=P\left(\frac{(101-1)S^2}{9}\le 133.33\right)-P\left(\frac{(101-1)S^2}{9}<100\right)\approx 0.99-0.90=0.09$$

### (٩,٧,٢) توزيع المعاينة لمتوسط (لنسبة) عينة من مجتمع برنولي

سنفترض على مدى هذه الفقرة أنّه لدينا مجتمع $[\Omega,\mathbb{S},\mathcal{P}]$ موصوف من خلال متغيّر عشـوائي $X$ خاضع لتوزيع برنولي بمَعْلَمة $0<p<1$ (احتمال النجاح)، ولنأخذ $\mathfrak{X}_n=\left(X_1,X_2,\ldots,X_n\right)$ عيّنة مسحوبة من هذا المجتمع. لقد ذكرنا في الفصل السابق أنّه في حالة المجتمعات البرنولية سنستخدم الرمز $\bar{\pi}_n$ بدلاً من $\bar{X}_n$ للتعبير عن متوسط العيّنة (انظر العلاقة [8-11-b])، وللأسباب التي ذكرناها سابقاً بشأن استخدام الرمز $\bar{X}$ بدلاً من $\bar{X}_n$ فإنّنا سنستخدم الرمز $\bar{\pi}$ بدلاً من $\bar{\pi}_n$ (والذي هو $\bar{X}_n$ نفسه). عندئذ بسبب وجود مَعْلَمة واحدة لهذا المجتمع فإنَّ الاهتمام ينصبّ على متوسط العيّنة $\bar{\pi}$ الذي يمثّل التكرار النسبي للنجاحات التي

نحصل عليها من هذه العيّنة؛ وذلك لأنَّ الملحوظات $X_i$ التي يوافقها نجاحاً ستأخذ القيمة 1 ، وأمّا باقي الملحوظات فسوف تأخذ القيمة 0 ، أي إنَّه لدينا في هذه الحالة:

$$\overline{\pi} = \frac{\text{عدد النجاحات في العينة}}{\text{حجم العينة}} \qquad [9,20]$$

لقد لاحظنا إنَّ مبرهنة النهاية المركزية لـ موافير–لابلاس تفيدنا في هذه المسألة بأنَّه سيكون المتغيِّر العشوائي $Z := \frac{\overline{\pi} - \mathbf{E}\,\overline{\pi}}{\sqrt{\mathbf{var}\,\overline{\pi}}}$ توزيع طبيعي معيّاري، ومن ثمَّ لو رمزنا لقيمة $\overline{\pi}$ بالرمز $\overline{p}$ فإنَّه سينتج لدينا الآتي:

أ- إذا كان حجم المجتمع غير محدود، وحجم العيّنة $n$ كبيراً بقدر كاف، فإنَّه سيكون لـ $\overline{\pi}$ توزيع طبيعي على وجه التقريب (انظر العلاقتين [8–17]) مع:

$$\mu_{\overline{\pi}} = p \qquad [9,21]$$

$$\sigma_{\overline{\pi}} = \sqrt{\frac{p\,(1-p)}{n}} \qquad [9,22]$$

ب- إذا كان حجم المجتمع المدروس محدوداً ويساوي $N$ ، وكان حجم العيّنة كبيراً نسبياً بحيث يكون $0.05N < n$ ، والسحب يتم دون إعادة، فإنَّه سيكون لـ $\overline{\pi}$ توزيع طبيعي على وجه التقريب بمتوسط $\mu_{\overline{\pi}} = p$ (كما في العلاقة [9,21]) ولكن بانحراف معيّاري:

$$\sigma_{\overline{\pi}} = \sqrt{\frac{p\,(1-p)}{n}}\sqrt{\frac{N-n}{N-1}} \qquad [9,22]$$

وتكون هذه المقادير أكثر دقّة كلما ازداد حجم العيّنة $n$ أكثر فأكثر، ويُلاحظ هنا كذلك أنَّه من أجل حجم عيّنة $n$ مثبَّت فإنَّ:

$$\lim_{N\to\infty} \sqrt{\frac{p\,(1-p)}{n}}\sqrt{\frac{N-n}{N-1}} = \sqrt{\frac{p\,(1-p)}{n}}\lim_{N\to\infty}\sqrt{\frac{N-n}{N-1}} = \sqrt{\frac{p\,(1-p)}{n}}$$

وتتوافق هذه الصيغة مع الحالة التي يكون فيها حجم المجتمع غير محدود أيضاً.

**(٩,٧,٢,١) ملاحظة**

نشير هنا إلى أنَّ النتائج السابقة تكون أكثر دقةً عندما يكون حجم العيّنة كبيراً بقدر كاف، وتتحسن النتائج أكثر فأكثر كلما ازداد حجم العينية أكثر فأكثر، وعلى أن يكون $5 \le n \cdot p$ وكذلك $5 \le n \cdot (1-p)$ أيضاً.

**(٩,٧,٢,٢) مثال**

ادعى مصنع لإنتاج القطع الإلكترونية أنَّ نسبة القطع المعيبة (غير مطابقة للمواصفات الفنية) في إنتاجه لا تتجاوز 0.05، وللتحقّق من صحة هذا الادعاء أخذنا عيّنة من الإنتاج الكلي لهذا المصنع بحجم $n = 225$ قطعة، فوجدنا بعد فحصها أنَّه يوجد في هذه العيّنة 15 قطعة معيبة. عندئذ:

أ- لنحسب احتمال أن يكون التكرار النسبي للقطع المعيبة في إنتاج المصنع أكبر من 0.05 .

ب- لنحسب احتمال أن يكون التكرار النسبي للقطع المعيبة في إنتاج المصنع أكبر من 0.05 إذا كان مستودع الإنتاج الكلي للمصنع يحوي 10000 قطعة فقط.

الحـل: من أجل الطلب:

أ- إنَّ الاحتمال المطلوب هو $P\,(\overline{\pi} \ge 0.05)$ ، علماً أنَّ الإحصاءة $\overline{\pi}$ يُعبِّر عن نسبة القطع المعيبة (النجاح هو الحصول على قطعة

معيبة ) في العيّنة $\mathfrak{X}_{225}$ ، وبما أنّه لم يذكر شيء عن عدد القطع الموجودة في مستودع الإنتاج الكلي لهذا المصنع فإنّنا سنفترض في هذه الحالة أنَّ حجم مجتمع القطع المنتجة غير محدود، ومن ثمَّ بملاحظة أنَّ قيمة مَعْلَمة مجتمع القطع المنتجة $p = 0.05$ فإنّه سيكون لدينا:

$$P\left(\overline{\pi} \geq 0.05\right) = P\left(\frac{\overline{\pi} - 0.05}{\sqrt{(0.05)(0.95)}}\sqrt{225} \geq 0\right) = 1 - P\left(Z < 0\right) = 1 - \Phi\left(0\right) = 0.5$$

علماً أنّنا قد وضعنا:

$$Z = \frac{\overline{\pi} - 0.05}{\sqrt{(0.05)(0.95)}}\sqrt{225}$$

حيث سيكون $Z \sim \mathbf{N}(0,1)$ . إنَّ هذه القيمة للاحتمال تثير لدينا الشك في أن يكون التكرار النسبي للقطع المعيبة في إنتاج هذا المصنع أكبر من 0.05 حقاً.

ب- إنَّ الاحتمال المطلوب هو $P\left(\overline{\pi} \geq 0.05\right)$ ، ولكن في هذه الحالة حجم مجتمع القطع محدولأنّه ذُكِرَ أنَّ مستودع الإنتاج الكلي لهذا المصنع يحوي 10000 قطعة فقط، وكذلك لدينا قيمة مَعْلَمة مجتمع القطع المنتجة $p = 0.05$ ، فعندئذ سيكون الاحتمال المطلوب يساوي:

$$P\left(\overline{\pi} \geq 0.05\right) = P\left(\frac{\overline{\pi} - 0.05}{\sqrt{(0.05)(0.950}}\sqrt{225}\sqrt{\frac{10000-225}{10000-1}} \geq 0\right)$$

$$= 1 - P\left(\frac{\overline{\pi} - 0.05}{\sqrt{(0.05)(0.95)}}\sqrt{225}\sqrt{\frac{10000-225}{10000-1}} < 0\right)$$

ومنه بوضع:

$$Z := \frac{\overline{\pi} - 0.05}{\sqrt{(0.05)(0.95)}}\sqrt{225}\sqrt{\frac{10000-225}{10000-1}}$$

سيكون $Z \sim \mathbf{N}(0,1)$ ، ومن ثمَّ ينتج لدينا أنَّ:

$$P\left(\overline{\pi} \geq 0.05\right) = 1 - P\left(Z < 0\right) = 1 - \Phi\left(0\right) = 0.5$$

وهنا نجد أنَّ قيمة الاحتمال التي حصلنا عليها تثير لدينا الشك أيضاً في أن يكون التكرار النسبي للقطع المعيبة في إنتاج هذا المصنع أكبر من 0.05 حقاً، وبذلك يتمّ الحل.

## (٩,٨) توزيعات المعاينة لمجتمعين مستقلين

**Sampling Distributions for two Independent Populations**

سنفترض على مدى هذه الفقرة أنّه لدينا مجتمعين إحصائيين $[\Omega_1, \mathbb{S}_1, \mathcal{P}_1]$ و $[\Omega_2, \mathbb{S}_2, \mathcal{P}_2]$ مستقلّان عن بعضهما البعض (وهذا يعني أنَّ الجبور $\mathbb{S}_1$ و $\mathbb{S}_2$ مستقلان عن بعضهما البعض)، وموصوفان من خلال متغيِّرين عشوائيين $X$ و$Y$ على الترتيب، وسنأخذ $\mathfrak{X}_n = (X_1, X_2, ..., X_n)$ و $\mathcal{Y}_m = (Y_1, Y_2, ..., Y_m)$ عينتين عشوائيتين من المجتمعين $[\Omega_1, \mathbb{S}_1, \mathcal{P}_1]$ و $[\Omega_2, \mathbb{S}_2, \mathcal{P}_2]$ على الترتيب، ولنفترض أنَّ للعينة $\mathfrak{X}_n$ متوسط $\overline{X}$ وانحراف معيّاري $S_X$، وكذلك للعينة $\mathcal{Y}_m$ متوسط $\overline{Y}$ وانحراف معيّاري $S_Y$، علماً أنَّ $n$ و $m$ عددان صحيحان موجبان مثبّتان.

### (١ ,٨ ,٩) توزيعات المعاينة لفرق متوسطي عينتين ونسبة التباين لعينتين من مجتمعين طبيعيين مستقلَّين

إذا كان المجتمعان $[\Omega_1, \mathcal{S}_1, \mathcal{P}_1]$ و $[\Omega_2, \mathcal{S}_2, \mathcal{P}_2]$ طبيعيين، أي إنَّ $X \sim \mathbf{N}(\mu_1, \sigma_1)$ و $Y \sim \mathbf{N}(\mu_2, \sigma_2)$، فعندئذ ينصبّ الاهتمام في هذه الحالة على دراسة توزيعات المعاينة للفرق $\bar{X} - \bar{Y}$ وللنسبة $\dfrac{S_X^2}{S_Y^2}$.

#### (١ ,١ ,٨ ,٩) توزيعات المعاينة للفرق $\bar{X} - \bar{Y}$

تحت الفرضيات التي ذكرت آنفاً فإنَّه لدراسة توزيعات المعاينة للفرق $\bar{X} - \bar{Y}$ سنناقش الحالات الآتية:

١- إذا كان تباين المجتمعين $\sigma_1^2$ و $\sigma_2^2$ معلومين، أي إنَّ $\sigma_1^2 = \sigma_{1,o}^2$ و $\sigma_2^2 = \sigma_{2,o}^2$، فعندئذ يكون للإحصاءة:

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\dfrac{\sigma_{1,o}^2}{n} + \dfrac{\sigma_{2,o}^2}{m}}}$$

توزيع طبيعي معيّاري، ومن ثمَّ ينتج لدينا أنَّ:

$$\mu_{\bar{X}-\bar{Y}} = \mu_1 - \mu_2 \qquad [9,23]$$

$$\sigma_{\bar{X}-\bar{Y}} = \sqrt{\frac{\sigma_{1,o}^2}{n} + \frac{\sigma_{2,o}^2}{m}} \qquad [9,24]$$

٢- إذا كان تباين المجتمعين $\sigma_1^2$ و $\sigma_2^2$ مجهولين ولكنَّهما متساويين، أي إنَّ $\sigma_1^2 = \sigma_2^2 = \sigma^2 = ?$، فعندئذ يكون للمتغيِّر العشوائي:

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2 [(1/n) + (1/m)]}}$$

توزيع طبيعي معيّاري، ومن ثمَّ يكون لكل من المتغيِّرين العشوائيين:

$$\chi_1 = \frac{(n-1) S_X^2}{\sigma^2} \qquad \& \qquad \chi_2 = \frac{(m-1) S_Y^2}{\sigma^2}$$

توزيع كاي مربّع بدرجات حرية $n-1$ و $m-1$ على الترتيب، ويبرهن على أنَّ لمجموعهما $\chi_1 + \chi_2$ توزيع كاي مربّع بدرجات حرية $n+m-2$، وبالنتيجة يمكننا أن نكتب الآتي:

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\sigma^2 [(1/n) + (1/m)]}} \sqrt{\frac{\sigma^2 (n+m-2)}{(n-1) S_X^2 + (m-1) S_Y^2}} = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{(n-1) S_X^2 + (m-1) S_Y^2}} \sqrt{\frac{\sigma^2 (n+m-2)}{\sigma^2 [(1/n) + (1/m)]}}$$

ومنه يكون لدينا:

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{(n-1) S_X^2 + (m-1) S_Y^2}} \sqrt{\frac{n+m-2}{(1/n) + (1/m)}} \qquad [9,25]$$

فإذا أخذنا الإحصاءة:

$$S_p^2 := \frac{(n-1)S_X^2 + (m-1)S_Y^2}{n+m-2} \qquad [9,26]$$

والتي تُدعى **المقدِّر المشترك لتباين العينتين** $\mathcal{X}_n$ **و** $\mathcal{Y}_m$ Pooled Estimator of Variance، فإنَّه بالتعويض في العلاقة [9,25] يبرهن على أنَّه سيكون للمتغيِّر العشوائي الآتي:

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p \sqrt{(1/n) + (1/m)}}$$

توزيع ستودنت بعدد درجات حريّة يساوي $n+m-2$، ومن ثمَّ يكون للإحصاءة $\bar{X} - \bar{Y}$ متوسط $\mu_{\bar{X}-\bar{Y}} = \mu_1 - \mu_2$ (كما في العلاقة [9,23]) ولكن بانحراف معيّاري يساوي:

$$\sigma_{\bar{X}-\bar{Y}} = \sqrt{\frac{(n+m)\left[(n-1)S_X^2 + (m-1)S_Y^2\right]}{n \cdot m \cdot (n+m-2)}} \qquad [9,27]$$

٣- إذا كان تباين المجتمعين $\sigma_1^2$ و $\sigma_2^2$ مجهولين وغير متساويين (أي إن $\sigma_1^2 = ?$ و $\sigma_2^2 = ?$ وكذلك $\sigma_1^2 \neq \sigma_2^2$)، فعندئذ دراسة هذه الحالة ليست بسيطة وتنطوي على صعوبات، ولذلك لن نخوض في تفاصيل هذه الحالة ولن نقدِّم أمثلة وتمارين من أجل هذه الحالة أيضاً. لكن سيكون للمتغيِّر العشوائي الآتي:

$$T^* = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\left(S_X^2 / n\right) + \left(S_Y^2 / m\right)}} \qquad [9,28]$$

توزيع ستودنت على وجه التقريب بعدد درجات حريّة يساوي:

$$\mathbf{df} = \left\lceil \frac{\left[\left(s_X^2 / n\right) + \left(s_Y^2 / m\right)\right]^2}{\sqrt{\left[\frac{1}{(n-1)}\left(s_X^2 / n\right)^2\right] + \left[\frac{1}{(m-1)}\left(s_Y^2 / m\right)^2\right]}} \right\rceil$$

ويطلق على التقدير السّابق لعدد درجات الحرِّية اسم **تقريب ستيرثويت** Satterthwaite's Approximation نسبة إلى الإحصائي **ستيرثويت** Satterthwaite, F. E.، وكذلك يُعرف باسم **درجات الحرية المُجمَّعة** pooled degrees of freedom، كما أنَّها تعرف باسم **معادلة ويلش-ستيرثويت** Welch–Satterthwaite equation، وقد نشره في عمل عام **1946**، علماً أنَّ $\lceil x \rceil$ هو أصغر عدد صحيح أكبر من $x$.

### (٩,٨,١,٢) ملاحظات

نودّ أن نشير هنا إلى أنَّه من أجل مجتمعين غير طبيعيين $[\Omega_1, \mathcal{S}_1, \mathcal{P}_1]$ بمتوسط $\mu_1$ وانحراف معيّاري $\sigma_1$ وكذلك $[\Omega_2, \mathcal{S}_2, \mathcal{P}_2]$ بمتوسط $\mu_2$ وانحراف معيّاري $\sigma_2$ تبقى النتائج السابقة بخصوص توزيعات المعاينة للفرق $\bar{X} - \bar{Y}$ صحيحة إذا كان حجم كل من العينتين $\mathcal{X}_n$ و $\mathcal{Y}_m$ كبيراً بقدر كاف، وتتحسن النتائج كلما ازداد حجم أي من العينتين أكثر فأكثر. أمّا إذا كان حجم أي من العينتين $\mathcal{X}_n$ و $\mathcal{Y}_m$ صغيراً فإنَّ تلكَ النّتائج ستصبح غير مرضية إلاّ من أجل المجتمعات الطبيعية حيث تحافظ على جودتها.

### (٩،٨،١،٣) توزيع المعاينة للنسبة $\frac{S_X^2}{S_Y^2}$ :

تحت الفرضيات المقدَّمة سابقاً في (١-٨-٩) يبرهن على أنَّه سيكون للمتغيِّر العشوائي الآتي:

$$F := \frac{S_X^2 \cdot \sigma_2^2}{S_Y^2 \cdot \sigma_1^2} \qquad [9,29]$$

توزيع فيشر بـ $m-1$ درجة حريَّة بسط و $n-1$ درجة حريَّة مَقام. أي إنَّ $F \sim \mathbf{F}(m-1, n-1)$، ومن ثمَّ بحسب المميِّزات العددية لمتغيِّر عشوائي خاضع لتوزيع فيشر ذي $m-1$ درجة حريَّة بسط و $n-1$ درجة حريَّة مَقام سيكون للإحصاءة $\frac{S_X^2}{S_Y^2}$ متوسط وتباين كما في العلاقتين الآتيتين على الترتيب:

$$\mu_{\frac{S_X^2}{S_Y^2}} = \frac{(n-1)\,\sigma_1^2}{(n-3)\,\sigma_2^2} \quad ; n \geq 4 \qquad [9,30]$$

$$\sigma^2_{\frac{S_X^2}{S_Y^2}} = \frac{2(n-1)^2(m+n-4)\sigma_1^4}{(m-1)(n-3)^2(n-5)\sigma_2^4} \quad ; m \geq 2,\ n \geq 6 \qquad [9,31]$$

نلاحظ هنا أنَّه إذا كان تباين المجتمعين متساويين (أي إنَّ $\sigma_1^2 = \sigma_2^2 = \sigma^2$ ) فعندئذ يبرهن على أنَّه سيكون للإحصاءة $F := \frac{S_X^2}{S_Y^2}$ توزيع فيشر ذو الـ $m-1$ درجة حريَّة و $n-1$ درجة حريَّة ولكن بمتوسط وتباين كما في العلاقتين الآتيتين على الترتيب:

$$\mu_{\frac{S_X^2}{S_Y^2}} = \frac{(n-1)}{(n-3)} \quad ; n \geq 4 \qquad [9,32]$$

$$\sigma^2_{\frac{S_X^2}{S_Y^2}} = \frac{2(n-1)^2(m+n-4)}{(m-1)(n-3)^2(n-5)} \quad ; m \geq 2,\ n \geq 6 \qquad [9,33]$$

### (٩،٨،١،٤) مثال

في متجرين A وB مستقلين عن بعضهما يوجد نوع من الملابس بقياسات ونماذج (موديلات) مختلفة، وقد أفادنا أصحاب هذين المتجرين أنَّ متوسط أسعار القطع في المتجر الأول A يساوي 53 وحدة نقدية بانحراف معيّاري قدره 3 وحدات نقدية بينما متوسط أسعار القطع في المتجر الثاني B يساوي 55 وحدة نقدية بانحراف معيّاري قدره وحدتين نقديتين. قمنا بسحب عيّنة من ملابس المتجر A بحجم 25 قطعة فوجدنا أنَّ متوسط أسعار قطع هذه العيّنة يساوي 52 وحدة نقدية بانحراف معيّاري قدره 4 وحدات نقدية، وكذلك قمنا بسحب عيّنة أخرى من ملابس المتجر B بحجم 36 قطعة فوجدنا أنَّ متوسط أسعار قطع هذه العيّنة يساوي 54 وحدة نقدية بانحراف معيّاري قدره 3 وحدات نقدية، فإذا علمنا أنَّ أسعار القطع تتوزع طبيعياً على وجه التقريب، فعندئذ:

أ- لنحسب احتمال أن يكون الفرق بين متوسطي العينتين لا يتجاوز وحدتين نقديتين.

ب- لنحسب احتمال أن تكون نسبة تباين العيّنة الثانية إلى تباين العيّنة الأولى أصغر من الواحد.

الحـل: بما أنَّه لم يُحدَّد حجم المجتمع فسوف نفترض أنَّ حجم هذا المجتمع غير محدود، ومن ثمَّ يكون من أجل الطلب:

أ- لدينا من مُعطيات المسألة التباين لكل من المجتمعين معلوماً، ومن ثمّ يمكننا أن نكتب الآتي:

$$P\left(\left|\overline{X} - \overline{Y}\right| \leq 2\right) = P\left(-2 \leq \overline{X} - \overline{Y} \leq 2\right) = P\left(\overline{X} - \overline{Y} \leq 2\right) - P\left(\overline{X} - \overline{Y} < -2\right)$$

ولكن لدينا:

$$P\left(\overline{X}-\overline{Y}\le 2\right)=P\left(\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}\le\frac{2-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}\right)$$

$$=P\left(\frac{(\overline{X}-\overline{Y})-(50-53)}{\sqrt{\frac{9}{25}+\frac{4}{36}}}\le\frac{2-(50-53)}{\sqrt{\frac{9}{25}+\frac{4}{36}}}\right)$$

$$=P\left(\left((\overline{X}-\overline{Y})-(50-53)\right)\left(\frac{9}{25}+\frac{4}{36}\right)^{-\frac{1}{2}}\le 7.28\right)=P\left(Z\le 7.28\right)=\Phi\left(7.28\right)\approx 1$$

وكذلك لدينا:

$$P\left(\overline{X}-\overline{Y}<-2\right)=P\left(\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}<\frac{-2-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}\right)$$

$$=P\left(\frac{(\overline{X}-\overline{Y})-(50-53)}{\sqrt{\frac{9}{25}+\frac{4}{36}}}<\frac{-2-(50-53)}{\sqrt{\frac{9}{25}+\frac{4}{36}}}\right)$$

$$=P\left(\left((\overline{X}-\overline{Y})-(50-53)\right)\left(\frac{9}{25}+\frac{4}{36}\right)^{-\frac{1}{2}}<1.46\right)=\Phi\left(1.46\right)=0.9495$$

وذلك لأنَّ $\mathbf{N}(0,1)\sim\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}=:Z$، ومن ثمَّ ينتج لدينا أنَّ الاحتمال المطلوب هو:

$$P\left(\left|\overline{X}-\overline{Y}\right|\le 2\right)\approx 1-0.9495=0.0505$$

٢- إنَّ الاحتمال المطلوب هو:

$$P\left(\frac{S_Y^2}{S_X^2}<1\right)=P\left(\frac{S_Y^2}{S_X^2}\frac{\sigma_1^2}{\sigma_2^2}<\frac{\sigma_1^2}{\sigma_2^2}\right)=P\left(F<\frac{9}{4}\right)=P\left(F<2.25\right)$$

حيث وضعنا $F:=\frac{9S_Y^2}{4S_X^2}$ الذي له توزيع فيشر $\mathbf{F}(24,35)$، ولذلك نجد من جدول توزيع فيشر أنَّ:

$$P\left(\frac{S_Y^2}{S_X^2}<1\right)=P\left(F<2.25\right)\approx 0.975$$

وذلك أنَّه لدينا أقرب قيمة لــ 2.25 هي 2.14 التي تقابل القيمة المقابلة لـ $m=24$ و $n=30$ حيث لدينا $f_{0.975}^{[24;30]}=2.14$.

### (٩،٨،٢) توزيعات المعاينة للفرق بين متوسطي (نسبتي) عينتين من مجتمعين برنوليين مستقلّين

إذا كان المجتمعان $[\Omega_1, \mathcal{S}_1, \mathcal{P}_1]$ و $[\Omega_2, \mathcal{S}_2, \mathcal{P}_2]$ برنوليين، أي إنَّ $X \sim \mathbf{B}(p_1)$ و $Y \sim \mathbf{B}(p_2)$، فعندئذ ينصبّ الاهتمام في هذه الحالة على دراسة توزيعات المعاينة للفرق $\overline{\pi}_1 - \overline{\pi}_2$، علماً أنَّ $\overline{\pi}_1$ و $\overline{\pi}_2$ هو التكرار النسبي للنجاحات في العيّنة $\mathfrak{X}_n$ و $\mathcal{Y}_m$ على الترتيب. الآن، وتحت الفرضيات المذكورة آنفاً يبرهن على أنّه سيكون للإحصاءة:

$$W = \frac{(\overline{\pi}_1 - \overline{\pi}_2) - (p_1 - p_2)}{\sqrt{\dfrac{m \cdot [p_1(1-p_1)] + n \cdot [p_2(1-p_2)]}{n \cdot m}}} \qquad [9,34]$$

توزيع طبيعي معيّاري بمتوسط وانحراف معيّاري على الترتيب هما:

$$\mu_{\overline{\pi}_1 - \overline{\pi}_2} = p_1 - p_2 \qquad [9,35]$$

$$\sigma_{\overline{\pi}_1 - \overline{\pi}_2} = \sqrt{\frac{m\,[p_1(1-p_1)] + n\,[p_2(1-p_2)]}{n \cdot m}} \qquad [9,36]$$

لدى التطبيق يُستعاض عن $p_1$ و $p_2$ بالقيم $\overline{p}_1$ و $\overline{p}_2$ على الترتيب لأنّنا ننطلق من أنَّ قيم مَعَالم المجتمع مجهولة كلياً أو جزئياً على الأقل، وذلك لأنّه لو كانت مَعَالم المجتمع معلومة لما كان هناك حاجة لهذه الدراسة من أجل جمع المعلومات حول المَعَالم.

#### (١، ٢، ٨، ٩) مثال

يودّ أحد تجار الجملة استيراد نوع من الأطعمة المعلّبة (الكنسروة) من إحدى شركتين تعمل كلّ منهما بشكلٍ مستقلّ عن الأخرى، حيث ادعت الشركة الأولى A أنَّ نسبة المعيب في إنتاجها منخفض جداً مقارنةً مع بقية الشركات المصنّعة لهذا المنتَج، بينما ادعت الشركة الثانية B أنَّ نسبة المعيب في إنتاجها منخفض جداً مقارنةً مع بقية الشركات المصنّعة لهذا المنتج أيضاً. قام التاجر بأخذ عيّنة بحجم $n = 144$ علبة من إنتاج الشركة A وتمَّ اختبارها فوجد 5 علب معيبة، ثمَّ أخذ عيّنة أخرى من إنتاج الشركة B بحجم $m = 100$ وتمَّ اختبارها فوجد 4 علب معيبة.

الآن، قرَّر هذا التاجر بأنّه لن يفاضل بين هاتين الشركتين (لن يُفضِّل إحداهما على الأخرى) إذا كان الفرق بين نسب العلب المعيبة لإنتاج الشركتين أكبر من 0.01 باحتمال أصغر من الـ 0.5 ، فإذا علمت أنَّ التوزيع الاحتمالي لجودة المنتج له توزيع طبيعي على وجه التقريب، فهل سيضطر التاجر إلى المفاضلة بين هاتين الشركتين، وفي حال المفاضلة أي الشركتين سيختار؟

**الحـل**: نلاحظ أنَّ إنتاج الشركتين يمكن تمثيله بمجتمعين $[\Omega_1, \mathcal{S}_1, \mathcal{P}_1]$ و $[\Omega_2, \mathcal{S}_2, \mathcal{P}_2]$ برنوليين بمَعْلمة $p_1$ و $p_2$ على الترتيب، وسنفترض أنَّ عدد العُلب المنتجة من قبل الشـركتين كبير جداً (لأنّه لم يذكر شـيء حول حجم الإنتاج لهما). عندئذ سيكون للإحصاءة $\overline{\pi}_1 - \overline{\pi}_2$ توزيع طبيعي معيّاري (وفقاً لمبرهنة موافير-لابلاس) بمتوسط وانحراف معيّاري على الترتيب هما:

$$\mu_{\overline{\pi}_1 - \overline{\pi}_2} = p_1 - p_2$$

$$\sigma_{\overline{\pi}_1 - \overline{\pi}_2} = \sqrt{\frac{m\,[p_1(1-p_1)] + n\,[p_2(1-p_2)]}{n \cdot m}}$$

وبما أنَّ $p_1$ و $p_2$ مجهولتان فإنّه يُستعاض عنهما بـ $\overline{p}_1 = \dfrac{5}{144} = 0.0347$ و $\overline{p}_2 = \dfrac{4}{100} = 0.04$ على الترتيب، ومن ثمَّ يصبح لدينا:

$$\mu_{\overline{\pi}_1 - \overline{\pi}_2} = \overline{p}_1 - \overline{p}_2 = 0.0347 - 0.04 = -0.0053$$

$$\sigma_{\overline{\pi}_1 - \overline{\pi}_2} = \sqrt{\frac{m\,[\overline{p}_1(1-\overline{p}_1)] + n\,[\overline{p}_2(1-\overline{p}_2)]}{n \cdot m}} = 0.0249$$

وعندئذ يتوقف قرار التاجر على معرفة قيمة الاحتمال $P\left(\left|\overline{\pi}_1-\overline{\pi}_2\right|\geq 0.01\right)$، حيث لدينا:

$$\begin{aligned}P\left(\left|\overline{\pi}_1-\overline{\pi}_2\right|\geq 0.01\right)&=1-P\left(\left|\overline{\pi}_1-\overline{\pi}_2\right|<0.01\right)=1-P\left(-0.01<\overline{\pi}_1-\overline{\pi}_2<0.01\right)\\&=1-P\left(\frac{-0.01-\mu_{\overline{\pi}_1-\overline{\pi}_2}}{\sigma_{\overline{\pi}_1-\overline{\pi}_2}}<\frac{(\overline{\pi}_1-\overline{\pi}_2)-\mu_{\overline{\pi}_1-\overline{\pi}_2}}{\sigma_{\overline{\pi}_1-\overline{\pi}_2}}<\frac{0.01-\mu_{\overline{\pi}_1-\overline{\pi}_2}}{\sigma_{\overline{\pi}_1-\overline{\pi}_2}}\right)\\&=1-P\left(\frac{-0.01+0.0053}{0.0249}<\frac{(\overline{\pi}_1-\overline{\pi}_2)+0.0053}{0.0249}<\frac{0.01+0.0053}{0.0249}\right)\end{aligned}$$

وبوضع $Z=\dfrac{(\overline{\pi}_1-\overline{\pi}_2)+0.0053}{0.0249}$ يصبح لدينا:

$$\begin{aligned}P\left(\left|\overline{\pi}_1-\overline{\pi}_2\right|\geq 0.01\right)&=1-P\left(0.61<Z<-0.19\right)=1-\left[\Phi(0.61)-\Phi(-0.19)\right]\\&=1-(0.7291-0.4247)=1-0.3044=0.6956\end{aligned}$$

إذاً سيضطر التاجر إلى المفاضلة بين هاتين الشركتين، ويكون في هذه الحالة اختيار الشركة A هو الأفضل لأنَّ نسب العلب المعيبة لديها (بحسب نتيجة العينة) هو الأقل.

بعد هذا التقديم الموجز لبعض توزيعات المعاينة سنقدِّم إحدى الصفات المهمّة التي يمكن لإحصاءة ما التمتّع بها. إنَّ هذه الصفة تُعرف باسم صفة الكفاية للإحصاءة.

## (٩,٩) الإحصاءة الكافية

### Sufficiency Statistic

ليكن لدينا $\left[\Omega,\mathbb{S},\mathcal{P}=\{P_\theta;\theta\in\Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ دالَّة كثافته الاحتمالية $f_X(x;\theta)$، ولنأخذ $\mathcal{X}_n=(X_1,X_2,\ldots,X_n)$ عيِّنة من هذا المجتمع، ولنفترض أنَّ $\mathbf{X}$ هي السّاحة التي تتحوَّل عليها قيم المتَّجه $\mathcal{X}_n$.

الآن لنأخذ الإحصاءات الآتية:

$$\left.\begin{aligned}T_1&=u_1(\mathcal{X}_n)\\T_2&=u_2(\mathcal{X}_n)\\&\vdots\\T_n&=u_n(\mathcal{X}_n)\end{aligned}\right\}\qquad [9,37]$$

وسنفترض أنَّ للمتَّجه $\mathrm{T}_n=(T_1,T_2,\ldots,T_n)$ دالَّة كثافة احتمالية $g(\vec{t}_n;\theta)$، وأنَّ السّاحة التي تتحوَّل عليها قيم هذا المتَّجه هي $\mathbf{T}$، فإذا كان لجملة المعادلات [9,37] حل وحيد من الشكل:

$$\left.\begin{aligned}x_1&=w_1(\vec{t}_n;\theta)\\x_2&=w_2(\vec{t}_n;\theta)\\&\vdots\\x_n&=w_n(\vec{t}_n;\theta)\end{aligned}\right\}\qquad [9,38]$$

وكان الشرطان الآتيان مُحقَّقين أيضاً:

أ- يوجد تقابل واحد لواحد بين قيم $\mathfrak{X}_n$ في السّاحة $\mathbf{X}$ وقيم $\mathbb{T}_n$ في السّاحة $\mathbf{T}$ .

ب- من أجل كل $k \in \mathbb{N}_n$ لدينا الدوال $w_k$ مستمرَّة وتملك مشتقات جزئية أولى مستمرَّة في كلٍّ من $t_1$، $t_2$،...و $t_n \in \mathbf{T}$ .

فعندئذ ستكون العلاقة الآتية مُحقَّقة:

$$g\left(\vec{t}_n;\theta\right) = \left|\mathbf{J}\right| f_{\mathfrak{X}_n}\left(w_1(\vec{t}_n;\theta), w_2(\vec{t}_n;\theta), \ldots, w_n(\vec{t}_n;\theta)\right) = \left|\mathbf{J}\right| \cdot \prod_{k=1}^{n} f_{X_k}\left(w_k(\vec{t}_n;\theta)\right) \qquad [9,39]$$

علماً أنَّ $\vec{t} := (t_1, t_2, \ldots, t_n)$ ، وأمّا $\mathbf{J}$ فهي قيمة المعين اليعقوبي Jacobean Determinant، وتحسب بالعلاقة الآتية:

$$\mathbf{J} \triangleq \frac{\partial(x_1, x_2, \ldots, x_n)}{\partial(t_1, x_2, \ldots, t_n)} = \begin{vmatrix} \frac{\partial x_1}{\partial t_1} & \frac{\partial x_1}{\partial t_2} & \cdots & \frac{\partial x_1}{\partial t_n} \\ \vdots & & \vdots & \vdots \\ \frac{\partial x_n}{\partial t_1} & \frac{\partial x_n}{\partial t_2} & \cdots & \frac{\partial x_n}{\partial t_n} \end{vmatrix} \qquad [9,40]$$

ومن العلاقة [9,39] يمكننا حساب دالَّة الكثافة الاحتمالية لأيّة إحصاءة $T_k$ مع $k \in \mathbb{N}_n$ باستخدام التوزيعات الهامشية، ومن ثمَّ بفرض $g_k(t_k;\theta)$ هي دالَّة الكثافة الهامشية لـ $T_k$ لكل $k \in \mathbb{N}_n$ ، فإنَّ:

$$h\left(t_2, t_3, \ldots, t_n;\theta \middle| t_1\right) := \begin{cases} \dfrac{g(\vec{t}_n;\theta)}{g_1(t_1;\theta)} & for \;\; g_1(t_1;\theta) > 0 \\ 0 & for \;\; g_1(t_1;\theta) = 0 \end{cases} \qquad [9,41]$$

هي دالَّة الكثافة الشرطية للمتَّجه $\mathbb{T}_n$ عندما يكون $T_1 = t_1$ مُحقَّقاً، فإذا كانت الدالَّة $h$ غير متعلِّقة بالمَعْلمة $\theta$ مهما تكن الإحصاءات $T_2$، $T_3$،...و $T_n$، فعندئذ تكون $T_1$ إحصاءة كافية للمَعْلمة $\theta$ .

إنَّ المعنى الاحتمالي لما سبق هو أنَّ أية إحصاءة أخرى غير $T_1$ لن تُقدِّم معلومات إضافية حول $\theta$ بأكثر مما قدَّمته الإحصاءة $T_1$ .

### (١ , ٩ , ٩) تعريف (الإحصاءة الكافية Sufficient Statistic)

تحت الفرضيات المذكورة آنفاً في بداية هذه الفقرة ليكن $T = g(\mathfrak{X}_n)$ إحصاءة مُعطاة، فعندئذ يُقال عن $T$ إنَّها **إحصاءة كافية** للمَعْلمة $\theta$ إذا كانت أيّة إحصاءة أخرى $U = h(\mathfrak{X}_n)$ لا تقدِّم معلومات إضافية حول المَعْلمة $\theta$ بأكثر مما قدَّمته الإحصاءة $T$ .

تجدر الإشارة هنا إلى أنَّ الإحصاءة الكافية لمَعْلمة $\theta$ ليست وحيدةً في الحالة العامّة، وكذلك التعامل مع تعريف الإحصاءة الكافية لتبيان إن كانت إحصاءة ما $T$ هي إحصاءة كافية لمَعْلمة المجتمع أم لا، هو أمر شاقّ في معظم الحالات، فلذلك تُقدَّم اختبارات سهلة التطبيق (بالمقارنة مع تطبيق التعريف) للحكم على إحصاءة ما إن كانت كافية لمَعْلمة المجتمع أم لا.

إنَّ مبرهنة فيشر- نويمان الآتية (التي سنقبلها دون برهان) تُعدّ أحد الاختبارات المستخدمة في إثبات إن كانت إحصاءة $T$ كافية لمَعْلمة المجتمع أم لا.

### (٢ , ٩ , ٩) مبرهنة (فيشر- نويمان Fisher -Neumann Proposition)

تحت الفرضيات التي ذكرت في بداية الفقرة (٩ , ٩)، وبفرض أنَّ $T_k = u_k(\mathfrak{X}_n)$ مع $k \in \mathbb{N}_n$ إحصاءات مُعطاة، فعندئذ تكون الإحصاءة $T_1$ كافية للمَعْلمة $\theta$ (**وبالمثل ينطبق القول على أي من الإحصاءات الأخرى**) **إذا وفقط إذا** كان ما يلي مُحقَّقاً:

$$f_{\mathfrak{X}_n}(\vec{x}_n;\theta) = w_1(u_1(\vec{x}_n);\theta) \cdot h(\vec{x}_n) \qquad [9,42]$$

علماً أنَّ $w_1(u_1(\vec{x}_n);\theta)$ هي دالَّة الكثافة الاحتمالية للإحصاءة $T_1$ ، وأمّا $h(\vec{x}_n)$ فهي دالَّة مستقلَّة عن $\theta$ من أجل أية قيمة $t_1 = u_1(\vec{x}_n)$ للإحصاءة $T_1$ .

تُقدِّم لنا المبرهنة الآتية (التي سنقبلها دون برهان أيضاً) اختباراً آخر لتبيان إن كانت إحصاءة ما $T$ كافية لمَعْلمة المجتمع $\theta$ أم لا.

**(٩,٩,٣) مبرهنة (التحليل إلى عوامل Analysis to Factors)**

تحت الفرضيات المذكورة في بداية الفقرة (٩,٩)، وبفرض أنَّ $T_k = u_k(\mathbb{X}_n)$ مع $k \in \mathbb{N}_n$ إحصاءات مُعطاة، فعندئذ تكون الإحصاءة $T_1$ كافية للمَعْلمة $\theta$ **إذا وفقط إذا** أمكننا أنَّ نكتب ما يلي:

$$\prod_{k=1}^{n} f_{X_k}(x_k;\theta) = \tilde{g}\left(u(\vec{x}_n);\theta\right)\cdot\tilde{h}(\vec{x}_n) \qquad [9,43]$$

علماً أنَّ $\tilde{h}$ دالَّة مستقلَّة عن $\theta$ و $\tilde{g}$ دالَّة متعلِّقة بالمَعْلمة $\theta$ في الحالة العامّة.

**(٩,٩,٣,١) أمثلة**

١- ليكن $[\Omega,\mathbb{S},\mathcal{P}]$ مجتمعاً برنولياً مَعْلمته $0<p<1$، ولنأخذ $\mathfrak{X}_n = (X_1, X_2, \ldots, X_n)$ عيّنة من هذا المجتمع، ولتكن $T$ إحصاءة تُمثِّل مجموع النجاحات في هذه العيّنة، أي إنَّ $T = u(\mathfrak{X}_n) = \sum_{k=1}^{n} X_k$، ولنبيِّن إن كانت هذه الإحصاءة كافية للمَعْلمة $p$ أم لا.

الحـل: نعلم أنَّ:

$$f_X(x;p) = p^x (1-p)^{1-x} \quad ; x \in \{0,1\}$$

ومن ثَّم يكون لدينا:

$$\prod_{k=1}^{n} f_{X_k}(x_k;p) = p^{\sum_{k=1}^{n} x_k} (1-p)^{n-\sum_{k=1}^{n} x_k} \qquad (1)$$

ولكن نعلم أنَّه إذا كانت $X_1$، $X_2$،...و $X_n$ متغيِّرات عشوائية برنولية مستقلَّة، فإنَّ مجموعها سيخضع للتوزيع الحدَّاني بمَعْلمتين $n$ و $p$، ومن ثَّم بوضع $t = \sum_{i=1}^{n} x_i$ فإنَّه سيكون لدالَّة الكثافة الاحتمالية للإحصاءة $T$ العرض الآتي:

$$w_T(t;p) = \binom{n}{t} p^t (1-p)^{n-t} \qquad (2)$$

و بسبب تحقّق العلاقتين (1) و (2) أنَّ نكتب الآتي:

$$\prod_{k=1}^{n} f_{X_k}(x_k;p) = \binom{n}{\sum_{k=1}^{n} x_k} p^{\sum_{k=1}^{n} x_k} (1-p)^{n-\sum_{k=1}^{n} x_k} \binom{n}{\sum_{k=1}^{n} x_k}^{-1}$$

ومنه بوضع:

$$h(\vec{x}_n) := \binom{n}{\sum_{k=1}^{n} x_k}^{-1} \qquad \& \qquad w_T\left(u(\vec{x}_n);p\right) = \binom{n}{\sum_{k=1}^{n} x_k} p^{\sum_{k=1}^{n} x_k} (1-p)^{n-\sum_{k=1}^{n} x_k}$$

تنتج لدينا العلاقة الآتية:

$$\prod_{k=1}^{n} f_{X_k}(x_k;p) = w_T\left(u(\vec{x}_n);p\right)\cdot h(\vec{x}_n)$$

ومنه بحسب اختبار فيشر–نويمان تكون $T$ إحصاءة كافية للمَعْلمة $p$.

٢- ليكن $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع للتوزيع الطبيعي $\mathbf{N}(\mu, \sigma)$ علماً أنَّ $\sigma$ معلومة، ولنأخذ $\mathcal{X}_n = (X_1, X_2, ..., X_n)$ عيِّنة من هذا المجتمع، ولنبيِّن فيما إذا كانت الإحصاءة $\overline{X}$ كافية لمتوسط المجتمع $\mu$ أم لا.

الحل: من أجل ذلك لدينا:

$$f_{\mathcal{X}_n}(\vec{x}_n; \mu) = \prod_{k=1}^{n} f_{X_k}(x_k; \mu) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k - \mu)^2\right]$$

$$= \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} \exp\left[-\frac{n(\overline{x} - \mu)^2 + \sum_{k=1}^{n}(x_k - \overline{x})^2}{2\sigma^2}\right]$$

ولكن من المعلوم أنَّ للإحصاء $\overline{X}$ توزيع $\mathbf{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$، ومن ثَمَّ يكون لدالَّة كثافته الاحتمالية العرض الآتي:

$$g_{\overline{X}}(u(\vec{x}_n); \mu) = \left(\frac{\sqrt{n}}{\sigma\sqrt{2\pi}}\right) \exp\left[-\frac{1}{2\sigma^2} n(\overline{x} - \mu)^2\right]$$

وبوضع:

$$\tilde{h}(\vec{x}_n) := \frac{1}{\sqrt{n}}\left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n-1}{2}} \exp\left[-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k - \overline{x})^2\right]$$

$$g_{\overline{X}}(u(\vec{x}_n); \mu) = \left(\frac{n}{2\pi\sigma^2}\right)^{\frac{1}{2}} \exp\left[-\frac{n(\overline{x} - \mu)^2}{2\sigma^2}\right]$$

فنجد أنَّ $\tilde{h}(\vec{x}_n)$ مستقلَّة عن المَعْلمة $\mu$، وكذلك لدينا:

$$f_{\mathcal{X}_n}(\vec{x}_n; \mu) = g_{\overline{X}}(u(\vec{x}_n); \mu) \cdot \tilde{h}(\vec{x}_n)$$

وهذا يعني بحسب اختبار فيشر– نويمان أنَّ الإحصاءة $\overline{X}$ هي إحصاءة كافية للمَعْلمة $\mu$.

تؤكد لنا المبرهنة الآتية (التي سنقبلها دون برهان أيضاً) أنَّ الإحصاءة الكافية لمَعْلمة مجتمع ليست وحيدةً في الحالة العامَّة.

**(٩، ٩، ٤) مبرهنة**

ليكن $[\Omega, \mathcal{S}, \mathcal{P} = \{P_\theta; \theta \in \Theta\}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ دالَّة كثافته الاحتمالية $f_X(x; \theta)$، ولنأخذ $\mathcal{X}_n$ عيِّنة من هذا المجتمع، ولنفترض أنَّ $T = g(\mathcal{X}_n)$ إحصاءة كافية للمَعْلمة $\theta$، وليكن $S = g(T)$ مع $g$ دالَّة تقابل، فعندئذ ستكون $S$ إحصاءة كافة للمَعْلمة $\theta$ أيضاً.

**(٩، ٩، ٥) أمثلة**

١- لقد بيَّنا في المثال / ١/ من (٩، ٩، ٣، ١) أنَّ الإحصاءة التي تُمثِّل مجموع النجاحات في عيِّنة $\mathcal{X}_n$ من مجتمع برنولي مَعْلمته $p$ هي إحصاءة كافية لهذه المَعْلمة، ومن ثَمَّ بحسب المبرهنة السابقة (٩، ٩، ٤) تكون الإحصاءة $\overline{\pi}$ التي تُمثِّل التكرار النسبي للنجاحات في العيِّنة $\mathcal{X}_n$ هي إحصاءة كافية للمَعْلمة $p$ أيضاً.

٢. لقد بيَّنا في المثال / ٢/ من (١ , ٣ , ٩ , ٩) أنَّ $\overline{X}$ هي إحصاءة كافية لـ $\mu$ متوسط مجتمع طبيعي مُعلَم تباينه، ومن ثمَّ بحسب المبرهنة السابقة (٤ , ٩ , ٩) ستكون الإحصاءة $a\overline{X}+b$ هي إحصاءة كافية لـ $\mu$ أيضاً من أجل أية قيم $a, b \in \mathbb{R}$ مع $a \neq 0$.

## (٩,١٠) كمية معلومات فيشر التي تقدمها عينة عشوائية

**Fisher's Information Quantity Provided by a Random Sample**

سنقدِّم فيما يلي كميّة معلومات فيشر التي يُقدِّمها عنصر من عيّنة، ومن ثمَّ الانتقال إلى كميّة معلومات فيشر التي تُقدِّمها عيّنة بحجم $n$. لكن تقديم هذا المفهوم يتطلب التعرف على بعض المفاهيم الأخرى التي سنوردها تباعاً.

**(١ , ١٠ , ٩) تعريف (المتغيِّر العشوائي النظامي Regular Random variables):**

ليكن لدينا $[\Omega, \mathbb{S}, \mathcal{P}=\{P_\theta; \theta\in\Theta\}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ دالَّة كثافته الاحتمالية $f_X(x;\theta)$، ولنأخذ $\mathfrak{X}_n=(X_1, X_2, ..., X_n)$ عيّنة من هذا المجتمع، فإذا كان من أجل كل $\theta \in \Theta_o \subseteq \Theta$ لدينا البندين الآتيين مُحقَّقين:

أ- القيم $\mathbf{E}\left(\frac{\partial \ln f_X(x;\theta)}{\partial\theta}\right)$ موجودة.

ب- لدينا:

$$\frac{\partial}{\partial\theta}\int_{-\infty}^{\infty} f_X(x;\theta)\,dx = 0 \quad \textit{for } X \textit{ continuous}$$

$$\frac{\partial}{\partial\theta}\sum_{i\in I} f_X(x_i;\theta) = 0 \quad \textit{for } X \textit{ discrete}$$

فعندئذ يُقال عن $X$ إنَّه **متغيِّر عشوائي نظامي** في المجموعة $\Theta_o$.

**(١ , ١ , ١٠ , ٩) مثال**

ليكن لدينا $[\Omega, \mathbb{S}, \mathcal{P}=\{P_\theta;\ \theta\in\Theta\}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع للتوزيع الأسّي بمَعْلَمة $\lambda>0$، فعندئذ نجد أنَّ هذا المتغيِّر العشوائي نظامي في أيّة فترة $(0,b)\subseteq(0,\infty)$ وذلك لأنَّه بفرض $\mathfrak{X}_n$ عيّنة مسحوبة من هذا المجتمع، فإنَّه سيكون لدينا:

$$\begin{aligned}
\mathbf{E}\left|\frac{\partial \ln f(x;\lambda)}{\partial\lambda}\right| &= \int_{-\infty}^{\infty}\left|\frac{\partial \ln f_X(x;\lambda)}{\partial\lambda}\right| f_X(x;\lambda)\,dx \\
&= \int_{-\infty}^{\infty}\left|\frac{1}{\lambda}-x\right|\lambda\,\mathbf{e}^{-\lambda x}\ \mathbf{I}_{[0,\infty)}(x)\,dx = \int_0^{\infty}\left(\frac{1}{\lambda}-x\right)\lambda\,\mathbf{e}^{-\lambda x}\,dx \\
&= \int_0^{\infty}\mathbf{e}^{-\lambda x}\,dx - \int_0^{\infty}\lambda x\,\mathbf{e}^{-\lambda x}\,dx = \frac{1}{\lambda}-\frac{1}{\lambda} = 0
\end{aligned}$$

وهذا يعني أنَّ $\mathbf{E}\left(\frac{\partial \ln f_X(x;\lambda)}{\partial\lambda}\right)=0$ موجودة.

وكذلك نجد:

$$\frac{\partial}{\partial\lambda}\int_{-\infty}^{\infty} f_X(x;\theta)\,dx = \frac{\partial}{\partial\lambda}\int_0^{\infty}\lambda\,\mathbf{e}^{-\lambda x}\,dx = 0$$

وهذا يعني أنَّ المتغيِّر العشوائي $X$ نظامي.

**(٩ , ١٠ , ١ , ٢) ملاحظات**

١– إذا كان $X$ متغيِّراً عشوائياً نظامياً، فعندئذ يُقال عن دالَّة توزيعه إنَّها دالَّة توزيع نظامية.

٢– إذا كان $X$ متغيِّراً عشوائياً نظامياً، وكان $T = g\left(\mathfrak{X}_n\right)$ إحصاءة كافية لـ $\theta$ ، فإنَّه يُبرهن على أنَّ $T$ نظامية أيضاً.

٣– إنَّ العيِّنة العشوائية المولَّدة من متغيِّر عشـوائي نظامي تُدعى **عيِّنة عشوائية نظامية**.

## (٩ , ١٠ , ٢) كميِّة معلومات فيشر التي يُقدِّمها عنصر من عيِّنة

**Fisher's Information Quantity Provided by an Element of a Sample**

ليكن $\left[\Omega, \mathcal{S}, \mathcal{P}=\{P_\theta ; \theta \in \Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ دالَّة كثافته الاحتمالية $f_X(x;\theta)$، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع. إنَّ **كميِّة معلومات فيشر** التي يُقدِّمها عنصر من هذه العيِّنة حول المَعْلمة $\theta$ سنرمز لها ب $\mathbf{I}(\theta)$ ، وتُعرَّف من خلال العلاقة الآتية:

$$\mathbf{I}(\theta) := \mathbf{E}\left(\frac{-\partial 2 \ \mathbf{ln} f_X \ (x \ ; \theta)}{\partial \theta^2}\right) \qquad [9,44]$$

**(٩ , ١٠ , ٢ , ١) ملاحظات**

١– يمكن إثبات أنَّ:

$$\mathbf{I}(\theta) = \mathbf{E}\left(\frac{-\partial 2 \ \mathbf{ln} f_X \ (x \ ; \theta)}{\partial \theta^2}\right) = \mathbf{E}\left(\frac{\partial \ \mathbf{ln} f_X \ (x \ ; \theta)}{\partial \theta}\right)^2 \geq 0 \qquad [9,45]$$

أي إنَّ كميِّة المعلومات التي يُقدِّمها عنصر من عيِّنة حول المَعْلمة $\theta$ هي مقدار غير سالب.

٢– إذا كان $\mathbf{I}(\theta) = 0$ فإنَّ ذلك يعني أنَّ $f_X(x;\theta)$ مستقلَّة عن $\theta$ ، وهذا يعني بدوره أنَّ المتغيِّر العشوائي $X$ **لا** يُقدم أية معلومات حول المَعْلمة $\theta$ ، ومن ثمَّ يمكن إثبات أنَّ دالَّة الكثافة $f_X(x;\theta)$ ثابتة باحتمال يساوي الواحد (بالنسبة إلى القياس $P_\theta$ ).

٣– سـنرمز لكمية المعلومات التي تقدِّمها إحصاءة $T = g\left(\mathfrak{X}_n\right)$ حول مَعْلَمة $\theta$ بالرمز $\mathbf{I}_T(\theta)$ .

تبيِّن لنا المبرهنة الآتية أنَّ كميِّة المعلومات التي تُقدِّمها عيِّنة تساوي كميِّة المعلومات التي يُقدِّمها عنصر من هذه العيِّنة مضروباً بحجم هذه العيِّنة.

**(٩ , ١٠ , ٢ , ٢) مبرهنة**

ليكن لدينا $\left[\Omega, \mathcal{S}, \mathcal{P}=\{P_\theta \ ; \ \theta \in \Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ دالَّة كثافته الاحتمالية $f_X(x;\theta)$، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، وبفرض أنَّ $\mathbf{I}_{\mathfrak{X}_n}(\theta)$ هي كميِّة المعلومات التي تُقدِّمها العيِّنة العشوائية $\mathfrak{X}_n$ ، فعندئذ ستكون العلاقة الآتية مُحقَّقة:

$$\mathbf{I}_{\mathfrak{X}_n}(\theta) = n \cdot \mathbf{I}(\theta) \qquad [9,46]$$

**البرهان:** بسبب الاستقلال والتطابق في التوزيع لملحوظات العيِّنة $X_1$، $X_2$، ...و $X_n$ يمكننا أن نكتب من أجل أي $\vec{x}_n = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ ما يلي:

$$\frac{\partial}{\partial \ \theta} \mathbf{ln} f_{\mathfrak{X}_n} \ (\vec{x}_n;\theta) = \frac{\partial}{\partial \theta} \mathbf{ln} \prod_{k=1}^{n} f_{X_k} \ (x_k;\theta) = \frac{\partial}{\partial \theta} \sum_{k=1}^{n} \mathbf{ln} f_{X_k} \ (x_k;\theta) = \sum_{k=1}^{n} \frac{\partial}{\partial \theta} \mathbf{ln} f_{X_k} (x_k;\theta)$$

وبالاشتقاق مرَّة أخرى بالنسبة إلى $\theta$ نجد الآتي:

$$\frac{\partial 2}{\partial \theta^2} \ln f_{\mathfrak{X}_n}(\vec{x}_n;\theta) = \sum_{k=1}^{n} \frac{\partial 2}{\partial \theta^2} \ln f_{X_k}(x_k;\theta)$$

ومن ثمَّ ينتج لدينا أنَّ:

$$\mathbf{I}_{\mathbb{X}_n}(\theta) = \mathbf{E}\left[-\frac{\partial 2}{\partial \theta^2} \ln f_{X_n}(\vec{x}_n;\theta)\right] = \sum_{k=1}^{n} \mathbf{E}\left[-\frac{\partial 2}{\partial \theta^2} \ln f_{X_k}(x_k;\theta)\right] = \sum_{k=1}^{n} \mathbf{I}(\theta) = n \cdot \mathbf{I}(\theta)$$

وبهذا يتمُّ البرهان.

**(٩,١٠,٢,٣) ملاحظة**

بفرض $\mathbf{I}_T(\theta)$ هي كميّة المعلومات التي تُقدِّمها إحصاءة $T = u(\mathfrak{X}_n)$ حول المَعْلمة $\theta$، فعندئذ يمكن التحقُّق من أنَّ:

$$\mathbf{I}_{\mathfrak{X}_n}(\theta) \geq \mathbf{I}_T(\theta)$$ [9,47]

حيث تتحقَّق إشارة المساواة عندما تكون $T$ إحصاءة كافية للمَعْلمة $\theta$.

**(٩,١٠,٢,٤) مثال**

ليكن لدينا $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع للتوزيع الطبيعي بمتوسط $\mu$ وانحراف معيّاري $\sigma$، ولتكن $\mathfrak{X}_n$ عيّنة من هذا المجتمع، ولنقم بحساب كميّة المعلومات التي تُقدِّمها هذه العيّنة حول المَعْلمة $\mu$، وبعد ذلك حول المَعْلمة $\sigma$.

**الحـل:** لدينا من أجل أي $\vec{x}_n = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$ ما يلي:

$$f_{\mathfrak{X}_n}(\vec{x}_n;(\mu,\sigma)) = \prod_{k=1}^{n} f_{X_k}(x_k;(\mu,\sigma)) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left[-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-\mu)^2\right]$$

ومن ثمَّ يكون لدينا:

$$\ln f_{\mathfrak{X}_n}(\vec{x}_n;(\mu,\sigma)) = n\left(-\ln\sigma - \ln\sqrt{2\pi}\right) - \frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-\mu)^2$$

والتي نجد منها أنَّ:

$$\frac{\partial 2 \ln f_{\mathfrak{X}_n}(\vec{x}_n;(\mu,\sigma))}{\partial \mu^2} = \frac{-n}{\sigma^2}$$

$$\frac{\partial 2 \ln f_{\mathfrak{X}_n}(\vec{x}_n;(\mu,\sigma))}{\partial \sigma^2} = \frac{n}{\sigma^2} - \frac{3}{\sigma^4}\sum_{k=1}^{n}(x_k-\mu)^2$$

ومن ثمَّ نجد أنَّ كميّة المعلومات التي تُقدِّمها العيّنة $\mathfrak{X}_n$ حول المَعْلمة $\mu$ تساوي:

$$\mathbf{I}_{\mathfrak{X}_n}(\mu) = \mathbf{E}\left(\frac{n}{\sigma^2}\right) = \frac{n}{\sigma^2}$$

وأمّا كميّة المعلومات التي تُقدِّمها هذه العيّنة حول المَعْلمة $\sigma$ فنجدها تساوي:

$$\begin{aligned}\mathbf{I}_{\mathfrak{X}_n}(\sigma) &= \mathbf{E}\left[\frac{3}{\sigma^4}\sum_{k=1}^{n}(X_k-\mu)^2-\frac{n}{\sigma^2}\right]=\frac{3}{\sigma^4}\sum_{k=1}^{n}\mathbf{E}(X_k-\mu)^2-\frac{n}{\sigma^2}\\ &= \frac{3}{\sigma^4}\underbrace{\sum_{k=1}^{n}\mathbf{E}(X-\mathbf{E}X)^2}_{\sigma^2}-\frac{n}{\sigma^2}=\frac{2n}{\sigma^2}\end{aligned}$$

**هذا ما تيسَّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل**

# تمارين الفصل التاسع

١- ليكن $\left[\Omega, \mathbb{S}, \mathcal{P}=\{P_\theta ; \theta \in \Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ بمتوسط $\mu$ وانحراف معيّاري $\sigma$ ، ولتكن $\mathfrak{X}_n=\left(X_1, X_2, \ldots, X_n\right)$ عيّنة مسحوبة من هذا المجتمع. عندئذ:

أ- إذا أخذنا $a_1$ ، $a_2$ ، ... و $0<a_n$ مع $\sum_{k=1}^{n} a_k=1$ ، ووضعنا $\tilde{X}=\sum_{k=1}^{n} a_k\ X_k$ ، حيث نلاحظ أنَّ $\tilde{X}$ هو المتوسط الموزون للعينة $\mathfrak{X}_n$ ، فما هي قيم $\mathbf{E}\,\tilde{X}$ و $\mathbf{var}\,\tilde{X}$ ، ومن ثمَّ قارن قيمة $\mathbf{var}\,\tilde{X}$ مع قيمة $\mathbf{var}\,\overline{X}$ ؟

ب- إذا أخذنا $Y=\left(\frac{X_1+X_2+\ldots+X_n}{n}\right)^2$، فما هي قيمة $\mathbf{E}\,Y$ ؟

ج- إذا أخذنا $Z=\frac{X_1^2+X_2^2+\ldots+X_n^2}{n}$ ، فما هي قيمة $\mathbf{E}\,Z$ ؟

٢- قمنا بسحب عيّنة $\mathfrak{X}_{100}$ (حجمها مئة) لأشخاص عاملين في مؤسسة تعليمية (جامعة مثلاً)، ومن ثمَّ دوِّنت عدد ساعات العمل خلال نصف عام دراسي لكلّ فرد منهم، فوجدنا أنَّ متوسط عدد ساعات عملهم خلال نصف عام دراسي يساوي 1000 ساعة بانحراف معيّاري قدره 20 ساعة عمل، فلو افترضنا أنَّ عدد ساعات العمل في هذه المؤسسة يتوزع طبيعياً على وجه التقريب، فعندئذ المطلوب ما يلي:

أ- حساب $P\left(997 \leq \overline{X} \leq 1003\right)$ .

ب- حساب $P\left(\overline{X}>1010\right)$ .

ج - بفرض أنَّ المؤسسة أعلمتنا أنَّ الانحراف المعياري لعدد ساعات العمل لديها هو 15 ساعة عمل، فما هي قيمة الاحتمال $P\left(S^2<225\right)$ ؟

٣- قمنا باستيراد نوع من السلع من مصدرين A وB مستقلّين عن بعضهما، حيث ادعى المصدر A أنَّ نسبة السلع غير المطابقة للمواصفات في إنتاجه لا تتجاوز 0.015 بينما ادعى المصدر B أنَّ نسبة السلع غير المطابقة للمواصفات في إنتاجه لا تتجاوز 0.005. قمنا بسحب عيّنة من السلع المستوردة من المصدر A بحجم 225 سلعة فوجدنا 3 سلع غير مطابقة للمواصفات، وكذلك سحبنا عيّنة من السلع المستوردة من المصدر B بحجم 289 سلعة فوجدنا 3 سلع غير مطابقة للمواصفات أيضاً، والمطلوب:

أ - ما هو احتمال أن تكون نسبة السلع غير المطابقة للمواصفات في إنتاج المصدر A ما بين 0.02 و0.04؟

ب- ما هو احتمال أنَّ تكون نسبة السلع غير المطابقة للمواصفات في إنتاج المصدر B أكبر من A بمقدار 0.005؟

جـ- ما هو احتمال أنَّ تكون نسبة السلع غير المطابقة للمواصفات في إنتاج المصدرين متساوية؟

د- ما هو احتمال أنَّ يكون الفرق بين نسب السلع غير المطابقة للمواصفات في إنتاج المصنعين أصغر من 0.005؟

هـ- كم يجب أنَّ يكون حجم العيّنة المسحوبة من المصدر A حتى يصبح احتمال أنَّ تكون نسبة السلع غير المطابقة للمواصفات في إنتاج المصدر A تساوي 0.015 أكبر من 0.9999؟

٤- في مصنع يوجد خطين لإنتاج المكثفات الكيميائية ذات سعة $47\,\mu F$ ميكرو فاراد، فإذا علمت أنَّ الخطين يعملان بشكل مستقل كل منهما عن الآخر، وأنَّ نسبة القطع المعيبة في إنتاج الخط الأول **(I)** لا تتجاوز 0.05، وأنَّ نسبة القطع المعيبة في إنتاج الخط الثاني **(II)** لا تتجاوز 0.04، وقمنا بأخذ عيّنة بحجم $n = 81$ من إنتاج الخط **(I)** فوجدنا 4 قطع معيبة (معطَّلة)، ثمَّ أخذنا عيّنة أخرى من إنتاج الخط **(II)** بحجم $m = 100$ فوجدنا 5 قطع معيبة، فعندئذ:

أ- احسب احتمال أن تكون نسبة القطع المعيبة في إنتاج المصنع **(I)** أكبر من 0.05 .

ب- احسب احتمال أن تكون نسبة القطع المعيبة في إنتاج المصنع **(II)** ما بين 0.03 و0.05.

جـ- احسب احتمال أن يكون الفرق بين نسب القطع المعيبة لإنتاج المصنعين أكبر من 0.15.

٥- ليكن $[\Omega, \mathbb{S}, \mathcal{P} = \{P_\theta; \theta \in \Theta\}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ دالَّة كثافته الاحتمالية مُعطاة من خلال العلاقة الآتية:

$$f_X(x;\theta) := \theta\, x^{\theta-1}\, \mathbf{I}_{(0,1)}(x) \quad ; \theta > 0$$

ولتكن $\mathfrak{X}_n$ عيّنة من هذا المجتمع، ولنأخذ الإحصاءة $T = u(\mathfrak{X}_n) = \prod_{k=1}^{n} X_k$، والمطلوب بيِّن إن كانت هذه الإحصاءة إحصاء كافية للمَعْلمة $\theta$ أم لا؟

٦- ليكن لدينا $[\Omega, \mathbb{S}, \mathcal{P} = \{P_\theta; \theta \in \Theta\}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع لتوزيع برنولي بمَعْلمة $0 < p < 1$، ولتكن $\mathfrak{X}_n$ عيّنة من هذا المجتمع، والمطلوب حساب كميّة معلومات فيشر التي تقدِّمها العينة $\mathfrak{X}_n$ .

٧- ليكن لدينا $[\Omega, \mathbb{S}, \mathcal{P} = \{P_\theta; \theta \in \Theta\}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع لتوزيع بواسون بمَعْلمة $0 < \lambda$، ولتكن $\mathfrak{X}_n$ عيّنة من هذا المجتمع، والمطلوب حساب كميّة معلومات فيشر التي تقدِّمها العينة $\mathfrak{X}_n$ .

٨- ليكن لدينا $[\Omega, \mathbb{S}, \mathcal{P} = \{P_\theta; \theta \in \Theta\}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع للتوزيع الأسّي بمَعْلمة $0 < \lambda$، ولتكن $\mathfrak{X}_n$ عيّنة من هذا المجتمع، والمطلوب حساب كميّة معلومات فيشر التي تقدِّمها العينة $\mathfrak{X}_n$ .

# الفصل العاشر

## مدخل إلى نظرية التقدير
## AN INTRODUCTION TO ESTIMATION THEORY

### تمهيد

نذكِّر أولاً إلى أنَّه عندما نذكر كلمة عيّنة (أو عيّنات) فإنَّنا نقصد بذلك أنَّها عيّنة (أو عيّنات) عشوائية، ومن ثَمَّ فإنَّ البيانات التي تكوِّنها هي بيانات عشوائية حتماً وليست كيفيّةً كما كانت في الفصول الثلاث الأولى من هذا الكتاب. كذلك عندما نأخذ $\mathfrak{X}_n$ عيّنة عشوائية فإنَّنا نقصد بذلك أنَّها بسيطة بحجم يساوي $n$، وأكثر من ذلك عندما نذكر كلمة مجتمع فإنَّنا نقصد بذلك أنَّه مجتمع إحصائيّ.

لقد لاحظنا أنَّه من أجل $[\Omega, \mathbb{S}, \mathcal{P}=\{P_\theta;\theta\in\Theta\}]$ مجتمع مُعطى فإنَّ القياسات الاحتمالية $P_\theta$ تفرض علينا معرفة كل المعلومات عن المَعْلمة $\theta$ بحيث يصبح معه النموذج الرياضياتي النظري $[\Omega, \mathbb{S}, \mathcal{P}=\{P_\theta;\theta\in\Theta\}]$ مناسباً للمجتمع الذي قيد الدراسة، وبما أنَّ المَعْلمة $\theta$ يمكن لها أن تكون متَّجها $\theta=(\theta_1,\theta_2,\ldots,\theta_k)$ من $\mathbb{R}^k$ فإنَّ مُركِّبات هذا المتَّجه قد تكون مجهولة كليًّا أو جزئيًّا في الحالة العامّة، ولهذا السبب يجب علينا سحب عيِّنات من المجتمع من أجل الحصول على معلومات تساعدنا في تقدير قيم هذه المَعَالِم المجهولة، بمعنى أنَّ المسألة التي لدينا آلت إلى تقدير هذه المَعَالِم المجهولة من خلال سحب عيِّنات من المجتمع. إنَّ القطاع الذي يهتمّ بتقدير مَعَالِم المجتمع Population Parameters وتقديم الاختبارات المتعلِّقة بهذه المَعَالِم، ومن ثَمَّ اتخاذ القرارات المناسبة بشأن المسألة الإحصائية التي قيد الدراسة، يُعرف باسم **الإحصاء الاستدلالي** Inferential Statistics وفي بعض المراجع يُذكر باسم **الاستدلال الإحصائي** Statistical Inference، ويرتكز هذا الفرع من الإحصاء الرياضيّاتي على الموضوعين الآتيين:

- نظرية التقدير.
- اختبارات الفرضيات الإحصائية.

علماً أنَّ موضوع اختبارات الفرضيات الإحصائية يهتمّ بتقديم اختبارات حول فرضيات مُحَدَّدة تتعلَّق بالمجتمع المدروس أو بمَعَالمه المجهولة (سيخصَّص الفصل التالي لهذا الموضوع). أمّا نظرية التقدير فإنَّها تهتمّ بشكل أساسي بتعيين الإحصاءات التي تقدِّر قيم المَعَالِم المجهولة للمجتمع بالإضافة إلى تقديم بعض الصفات المميِّزة لهذه المقدِّرات، وكذلك تحديد فترات تنتمي إليها تلك المَعَالِم المجهولة باحتمال معلوم مستندة في ذلك على مقدِّرات تلك المَعَالِم المجهولة.

نشير هنا إلى أنَّنا لن نسترسل كثيراً في شرح فقرات هذا الفصل حيث يمكن لمن يرغب بالمزيد الرجوع إلى المراجع ذات الصلة ومنها [6]، [22]، [24]، [33]، [37]، [40]، [45]، [47]، [62] و[66] في فهرس مراجع هذا الكتاب، وأخيراً نشير إلى أنَّنا سنركِّز في دراستنا على تقدير المَعَالِم المجهولة لمجتمع وفقا للمسارات الآتية:

١- كيفية تعيين مُقدِّر مَعْلمة مجتمع.
٢- البحث في بعض الصفات المميِّزة للمقدِّرات.
٣- تقديم فترات تنتمي إليها المَعَالِم المجهولة للمجتمع باحتمال معلوم، وهذه الفترات تُدعى فترات الثّقة.

## (١٠,١) تعيين مقدر معلمة مجتمع إحصائي

**Determination an Estimator of Population Parameter**

نقوم فيما يلي بتقديم طريقتين فقط لـ تعيين مُقدِّر مَعْلمة مجتمع وهما:

١- **طريقة العزوم** Moments Method.

٢- **طريقة الأرجحية العظمى** Maximum Likelihood Method.

علماً أنَّه توجد طرائق أخرى تـ ستخدم في تعيين مُقدِّر مَعْلمة مجتمع حيث يمكن للقارئ الاطلاع على بعضها بالرجوع إلى المراجع ذات الصلة ومنها ذكر آنفاً.

### (١٠,١,١) طريقة العزوم في تعيين مُقدِّرات مَعَالم مجتمع

ليكن $[\Omega, \mathbb{S}, \mathcal{P} = \{P_\theta ; \theta \in \Theta\}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، وسنفترض أنَّ $\theta$ متَّجه بـ $k$ مُركَّبة، أي إنَّ $\theta = (\theta_1, \theta_2, ..., \theta_k)$، ولنأخذ $\mathfrak{X}_n = (X_1, X_2, ..., X_n)$ عيِّنة من هذا المجتمع. عندئذ تقوم طريقة العزوم في تعيين مُقدِّر مَعْلمة المجتمع $\theta$ على الأسس الآتية:

أ- حساب العزم الابتدائي من المرتبة $r$ لكل من $\mathfrak{X}_n$ و $X$ من أجل جميع القيم $r \in \mathbb{N}_k$.

ب- حل جملة المعادلات الآتية (حلاً مشتركاً):

$$\mathbf{M}^{(r)} = \mu^{(r)} |_{\theta = \hat{\theta}} \quad ; \forall r \in \mathbb{N}_k \qquad [10,1]$$

علماً أنَّ $\mathbf{M}^{(r)}$ و $\mu^{(r)}$ هما العزم الابتدائي من المرتبة $r$ لـ $\mathfrak{X}_n$ و $X$ على الترتيب.

ج- بفرض أنَّ $\hat{\theta}_1$، $\hat{\theta}_2$، ...و $\hat{\theta}_k$ هي الحلول النّاتجة عن جملة المعادلات [10,1]، فإنَّنا نضع $\hat{\theta} = (\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k)$، فعندئذ يُقال عن $\hat{\theta}$ إنَّه **مقدِّر المَعْلمة** $\theta$ وفقاً لطريقة العزوم، وأمّا الحلول $\hat{\theta}_1$، $\hat{\theta}_2$، ...و $\hat{\theta}_k$ فإنَّها تُدعى **المقدِّرات الإحصائية** Statistical Estimators (أو المقدِّرات) للمَعَالم $\theta_1$، $\theta_2$، ...و $\theta_k$ على الترتيب.

إنَّ هذه الطريقة في تقدير مَعْلمة مجتمع هي أول الطرائق التي قدِّمت لـ تعيين مُقدِّر مَعْلمة مجتمع، وعلى وجه التحديد تعود إلى الإحصائي الإنكليزي بيرسون.

#### (١٠,١,١,١) أمثلة

١- ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ له توزيع طبيعي $\mathbf{N}(\mu, \sigma)$، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، ولنقم بتعيين مُقدِّري المَعْلمتين $\mu$ و $\sigma$.

الحـل: من أجل تعيين مُقدِّري المَعْلمتين $\mu$ و $\sigma$ يجب علينا حل جملة المعادلات [10,1] الخاصَّة بهذه المسألة (هنا لدينا معادلتان فقط لأنَّه لدينا مَعْلمتان فقط)، وبما أنَّه لدينا معلمتين فإنَّه يتوجّب علينا تعيين المقادير الآتية:

$$\mathbf{M} = \frac{X_1 + X_2 + ... + X_n}{n} = \bar{X} \qquad \& \qquad \mathbf{M}^{(2)} = \frac{X_1^2 + X_2^2 + ... + X_n^2}{n} \triangleq \overline{X^2}$$

$$\mu = \mathbf{E}X = \mu \qquad \& \qquad \mu^{(2)} = \mathbf{E}X^2 = \sigma^2 + \mu^2$$

ومن ثمَّ يكون لدينا:

$$\mathbf{M} = \mu|_{(\mu,\sigma) = (\hat{\mu},\hat{\sigma})} \Rightarrow \frac{X_1 + X_2 + ... + X_n}{n} = \mu \Big|_{(\mu,\sigma) = (\hat{\mu},\hat{\sigma})} \Rightarrow \bar{X} = \hat{\mu} \qquad (1)$$

وكذلك:

$$\mathbf{M}^{(2)} = \mu^{(2)}\Big|_{(\mu,\sigma)=(\hat{\mu},\hat{\sigma})} \Rightarrow \frac{X_1^2 + X_2^2 + \ldots + X_n^2}{n} = \sigma^2 + \mu^2\Big|_{(\mu,\sigma)=(\hat{\mu},\hat{\sigma})} \Rightarrow \overline{X^2} = \hat{\sigma}^2 - \hat{\mu}^2 \qquad \textbf{(2)}$$

وبالتعويض في **(2)** عن $\hat{\mu}$ بما يساويها في العلاقة **(1)** ينتج لدينا أنَّ $\hat{\sigma}^2 = \overline{X^2} - \bar{X}^2$، ومن ثمَّ يكون $\bar{X}$ هو مُقدِّر المَعْلمة $\mu$، وأما مُقدِّر المَعْلمة $\sigma$ فهو $\hat{\sigma} = +\sqrt{\overline{X^2} - \bar{X}^2}$، وبالتالي فإنَّ المتَّجه:

$$\hat{\theta} = \left(\bar{X}\ ,\ \sqrt{\overline{X^2} - \bar{X}^2}\right)$$

هو مقدِّر المَعْلمة $\theta = (\mu,\sigma)$ وفقاً لطريقة العزوم.

٢- ليكن $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X \sim \mathbf{B}(n,p)$، ولنأخذ $\mathcal{X}_n$ عيِّنة من هذا المجتمع، فعندئذ لتعيين مُقدِّري المَعْلمتين $n$ و $p$ يجب علينا حل جملة المعادلات [10,1] الخاصَّة بهذه المسألة (**هنا لدينا معادلتان فقط لأنَّه لدينا مَعْلمتان فقط**)، وبما أنَّه لدينا معلمتين فإنَّه يتوجَّب علينا تعيين المقادير الآتية:

$$\mathbf{M} = \frac{X_1 + X_2 + \ldots + X_n}{n} = \bar{X} \qquad \& \qquad \mathbf{M}^{(2)} = \frac{X_1^2 + X_2^2 + \ldots + X_n^2}{n} \triangleq \overline{X^2}$$

$$\mu = \mathbf{E}X = n \cdot p \qquad \& \qquad \mu^{(2)} = \mathbf{E}X^2 = n\,p\,(1-p) + n^2 p^2$$

ومن ثمَّ يكون:

$$\mathbf{M} = \mu\,\Big|_{(n,p)=(\hat{n},\hat{p})} \Rightarrow \bar{X} = n \cdot p\,\Big|_{(n,p)=(\hat{n},\hat{p})} \Rightarrow \bar{X} = \hat{n} \cdot \hat{p} \qquad \textbf{(1)}$$

وكذلك لدينا:

$$\mathbf{M}^{(2)} = \mu^{(2)}\Big|_{(n,p)=(\hat{n},\hat{p})} \Rightarrow \overline{X^2} = n\,p\,(1-p) + n^2 p^2\Big|_{(n,p)=(\hat{n},\hat{p})} \Rightarrow \overline{X^2} = \hat{n}\,\hat{p}\,(1-\hat{p}) + \hat{n}^2\hat{p}^2 \qquad \textbf{(2)}$$

وبالحل المشترك للمعادلتين **(1)** و **(2)** نجد منهما الحلين $\hat{n} = \dfrac{1}{\bar{X} - \overline{X^2}}$ و $\hat{p} = \dfrac{\bar{X} - \overline{X^2}}{\bar{X}}$، والذين يقدِّران المَعْلمتين $n$ و$p$، وبالتالي فإنَّ المتَّجه:

$$\hat{\theta} = \left(\frac{1}{\bar{X} - \overline{X^2}}\ ,\ \frac{\bar{X} - \overline{X^2}}{\bar{X}}\right)$$

هو مقدِّر المَعْلمة $\theta = (n,p)$ وفقاً لطريقة العزوم حيث نلاحظ هنا أنَّه يجب أنَّ يكون $\overline{X^2} < \bar{X}$.

٣- ليكن $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X \sim \mathbf{P}(\lambda)$ (**أي إنه لدينا مجتمع بواسوني بمَعْلمة** $0 < \lambda$)، ولنأخذ $\mathcal{X}_n$ عيِّنة من هذا المجتمع، فعندئذ لتعيين مُقدِّر المَعْلمة $\lambda$ يجب علينا حل جملة المعادلات [10,1] الخاصَّة بهذه المسألة، ولكن بما أنَّه لدينا مَعْلمة واحدة فقط فإنَّه يكفي حل المعادلة $\mathbf{M} = \mu\big|_{\lambda=\hat{\lambda}}$ فقط، حيث لدينا:

$$\mathbf{M} = \bar{X} \qquad \& \qquad \mu = \mathbf{E}X = \lambda$$

ومن ثَمَّ نحصل على الحل $\hat{\lambda} = \bar{X}$، وبالتالي فإنَّ $\bar{X}$ هو مُقدِّر المَعْلمة $\lambda$ وفقاً لطريقة العزوم.

٤- بفرض أنَّ $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمع موصوفاً من خلال متغيِّر عشوائي $X \sim \mathbf{Ex}(\lambda)$ (**أي إنه لدينا مجتمع أسِّي بمَعْلمة** $0 < \lambda$)،

ولنأخذ $\mathfrak{X}_n$ عيّنة من هذا المجتمع، فعندئذ لتعيين مُقدِّر مَعْلمة هذا المجتمع (وفقاً لطريقة العزوم) يكفي حل المعادلة $\mathbf{M} = \mu\,|_{\lambda=\hat{\lambda}}$، حيث لدينا:

$$\mathbf{M} = \bar{X} \qquad \& \qquad \mu = \mathbf{E}X = \frac{1}{\lambda}$$

فنحصل نتيجة لذلك على الحل $\hat{\lambda} = \frac{1}{\bar{X}}$، ومن ثمَّ يكون $\hat{\lambda} = \frac{1}{\bar{X}}$ هو مُقدِّر المَعْلمة $\lambda$ وفقاً لطريقة العزوم.

### (٢ , ١ , ١ , ١٠) ملاحظات

١ - إنَّ استخدام هذه الطريقة غير ممكن من أجل المجتمعات التي لا تملك عزوماً ابتدائية (ومنها المجتمعات التي تتبع توزيع كوشي).

٢ - إنَّ المقدِّرات العزومية لمَعْلمة مجتمع ليست وحيدة في الحالة العامّة، إذ يمكن حسابها باستخدام العزوم المركزية أيضاً.

## (٢ , ١ , ١٠) طريقة الأرجحية العظمى في تعيين مُقدِّرات مَعَالم مجتمع

**Maximum Likelihood Method for Estimations of Population Parameters**

إنَّ طريقة الأرجحية العظمى (أو الاحتمال الأكبر) تُعدّ من أكثر الطرائق المستخدمة في تعيين مُقدِّرات مَعَالم المجتمعات لما تتمتَّع به من صفات مرغوبة في المقدِّرات (التي سنوردها فيما بعد)، ولكن قبل البدء بتقديم هذه الطريقة سنقدِّم مثالاً بسيطاً يوضِّح لنا مفهوم الأرجحية العظمى.

### (١ , ٢ , ١ , ١٠) مثال

في مدينة ما تبيَّن أنَّ نسبة الذكور والإناث هما 52% و48% ولكن دون معرفة أي منهما للذكور أو للإناث، ولنفترض أنَّ حصولنا على ذكر عند سحب فرد من هذا المجتمع يُعدّ نجاحاً. عندئذ بفرض أنَّ جميع عناصر هذا المجتمع مستقلّة بعضها عن البعض الآخر، فإنَّ كل عمليّة سحب لفرد من هذا المجتمع هي تجربة برنولية بمَعْلمة $p = ?$، فإذا كان $X$ متغيِّر عشوائي واصف لهذا المجتمع فإنَّه سيكون لـ $X$ توزيع برنولي بدالَّة كثافة احتمالية $f_X(x;p)$، أي إنَّه لدينا:

$$f_X(x;p) = P(X = x;p) = p^x\,(1-p)^{1-x}\, I_{\{0,1\}}(x)$$

علماً أنَّ $p$ ينتمي إلى المجموعة $\Theta = \{0.48, 0.52\} \subseteq \mathbb{R}$، ومن ثمَّ يكون لدينا ما يلي:

$$\text{If } \; p = 0.48 \Rightarrow f_X(0;p) = 0.52 \;\; \& \;\; f_X(1;p) = 0.48$$

$$\text{If } \; p = 0.52 \Rightarrow f_X(0;p) = 0.48 \;\; \& \;\; f_X(1;p) = 0.52$$

وهكذا نلاحظ أنَّه:

١ - إذا كان $x = 0$ فإنَّنا نرجِّح استخدام القيمة $\hat{p} = 0.48$ على القيمة $\hat{p} = 0.52$ كتقدير لـ $p$ لأنَّ للدالَّة $f_X(0;p)$ احتمال أكبر (ومن ثمَّ أرجحية أكبر) من أجل هذه الحالة.

٢ - إذا كان $x = 1$ فإنَّنا نرجِّح استخدام القيمة $\hat{p} = 0.52$ على القيمة $\hat{p} = 0.48$ كتقدير لـ $p$ لأنَّ للدالَّة $f_X(1;p)$ احتمال أكبر (ومن ثمَّ أرجحية أكبر) من أجل هذه الحالة.

الآن وبشكل مماثل يمكننا مناقشة الحالة التي توافق سحب عيّنة $\mathfrak{X}_n = (X_1, X_2, \ldots, X_n)$ من مجتمع، فعلى سبيل المثال لو عُدنا إلى المثال السابق، وقمنا بسحب عيّنة حجمها يساوي 5، أي لدينا $\mathfrak{X}_n = (X_1, X_2, X_3, X_4, X_5)$، فعندئذ يكون للمتغيِّر العشوائي الواصف لهذا المجتمع توزيع حدَّاني (لأنَّه لدينا تكرارات مستقلّة لتجارب برنولية)، ومن ثمَّ يكون لـ $f_X(\bar{x}_5;p)$ العرض الآتي:

$$f_X(x;p)=\sum_{x=0}^{5}\binom{5}{x}p^x(1-p)^{5-x}\,I_{\{0,1,2,3,4,5\}}(x)$$

وعلى سبيل التبسيط والتوضيح يمكننا عرض النتائج لهذه الدراسة من خلال الجدول الآتي:

الجدول (١٠ ,١)

| | $p$ | 0.48 | 0.52 |
|---|---|---|---|
| | $f_X(x;p)$ | $f_X(x;0.48)$ | $f_X(x;0.52)$ |
| $x$ | 0 | 0.0380204 | 0.0254804 |
| | 1 | 0.1754780 | 0.1380188 |
| | 2 | 0.3239608 | 0.2990408 |
| | 3 | 0.2990408 | 0.3239608 |
| | 4 | 0.1380188 | 0.1754788 |
| | 5 | 0.0254804 | 0.0380204 |

علماً أنَّنا استخدمنا الحاسبة الإلكترونية للوصول إلى قيم دقيقة للاحتمالات. حيث نلاحظ بوضوح أنَّه إذا كان $x=0$ أو $x=1$ أو $x=2$ فإنَّنا نرجِّح استخدام القيمة $\hat{p}=0.48$ على القيمة $\hat{p}=0.52$ كتقدير لـ $p$ لأنَّه للدالَّة $f_X(x;p)$ احتمال أكبر (ومن ثمَّ أرجحية أكبر) من أجل هذه القيم لـ $x$ ، وأمّا إذا كان $x=3$ أو $x=4$ أو $x=5$ فإنَّنا نرجِّح استخدام القيمة $\hat{p}=0.52$ على القيمة $\hat{p}=0.48$ كتقدير لـ $p$ لأنَّه للدالَّة $f_X(x;p)$ احتمال أكبر من أجل هذه الحالات الثلاث الأخيرة.

### (١٠ ,١ ,٢ ,٢) تعريف (دالَّة الأرجحية لعيِّنة Likelihood Function of a sample)

ليكن لدينا $\left[\Omega,\mathcal{S},\mathcal{P}=\{P_\theta;\theta\in\Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ ، ولتكن $\mathcal{X}_n=(X_1,X_2,\dots,X_n)$ عيِّنة من هذا المجتمع، ولنضع بالتعريف ما يلي:

$$L(\vec{x}_n;\theta)=f_{\mathrm{X}_n}(\vec{x}_n;\theta)=\prod_{k=1}^{n}f_X(x_k;\theta) \qquad [10,2]$$

علماً أنَّ $\vec{x}_n$ هي القيمة الملاحظة للعيِّنة العشوائية $\mathcal{X}_n$ ، فعندئذٍ يُطلق على الدالَّة $L(\vec{x}_n;\theta)$ اسم **دالَّة الأرجحية** للعيِّنة $\mathcal{X}_n$ .

بالرغم من أنَّ العديد من المؤلَّفات استخدمت تسمية **دالَّة الأرجحية** إلاَّ أنَّ مؤلَّفات عربية أخرى استخدمت تسميات أخرى بدلاً من عبارة دالَّة الأرجحية، ومن هذه التسميات: **دالَّة الجوازيَّة** أو **دالَّة المعقولية** أو **دالَّة الإمكان**، وقد اخترنا التسمية المذكورة آنفًا لاعتقادنا بأرجحية معناها مع التفسير العلمي لتلك الدالَّة.

### (١٠ ,١ ,٢ ,٣) تعريف (مُقدِّر الأرجحية العظمى Maximum Likelihood Estimator)

ليكن لدينا $\left[\Omega,\mathcal{S},\mathcal{P}=\{P_\theta;\theta\in\Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ ، ولتكن $\mathcal{X}_n=(X_1,X_2,\dots,X_n)$ عيِّنة من هذا المجتمع، وبفرض أنَّ $L(\vec{x}_n;\theta)$ هي دالَّة الأرجحية للعيِّنة $\mathcal{X}_n$ ، فعندئذٍ يُطلق على الحل $\hat{\theta}$ الذي من أجله يكون لدينا:

$$L(\vec{x}_n;\hat{\theta})\ge L(\vec{x}_n;\theta) \qquad for\ all\ \ \theta\ne\hat{\theta}$$

اسم **مُقدِّر الأرجحية العظمى** للمَعْلمة $\theta$ ، حيث يُستعاض عن $\vec{x}_n$ بـ $\mathcal{X}_n$ في الصيغة النهائية لـ $\hat{\theta}$ .

إنَّ هذا المبدأ في التقدير يعني أنَّنا سنختار من مُقدِّرات المَعْلمة $\theta$ تلك المقدِّرات التي يكون من أجلها للعيِّنة $\mathcal{X}_n$ أكبر قيمة احتمالية، وهذا يعني أنَّه يمكننا أن نصيغ العلاقة الأخيرة على النحو الآتي:

$$L\ (\vec{x}_n\ ;\ \hat{\theta}) = \max_{\theta\in\Theta}\left\{ L\,(\vec{x}_n\ ;\ \theta)\ \big|_{\theta=\hat{\theta}}\right\} \qquad [10,3]$$

### (٤, ٢, ١, ١٠) ملاحظات

١- تبدي هذه الطريقة فعالية كبيرة مع التوزيعات التي لها شكل دوال أسيَّة.

٢- من الحالات المهمّة أيضاً عند تعيين مُقدِّرات الأرجحية العظمى، تلك الحالة التي تكون فيها دالَّة الأرجحية $L\,(\vec{x}_n;\theta)$ قابلة للاشتقاق الجزئي بالنسبة إلى $\theta$، إذ إنَّنا نعلم أنَّ الدالة $L\,(\vec{x}_n;\theta)$ تأخذ قيمة عظمى عندما يكون $\frac{\partial L(\vec{x}_n\,;\,\theta)}{\partial\,\theta}=0$، ومن ثمّ يمكن تعيين مُقدِّر الأرجحية العظمى $\hat{\theta}$ للمَعْلمة $\theta$ من خلال العلاقة الآتية:

$$\frac{\partial\ L(\vec{x}_n\,;\,\theta)}{\partial\ \theta}\Bigg|_{\theta=\hat{\theta}} = 0 \qquad [10,4]$$

وفي حال كان لدينا $\theta=(\theta_1,\theta_2,...,\theta_k)$ فإنَّه من أجل تطبيق هذه الملاحظة يجب أن تكون الدالَّة $L\,(\vec{x}_n;\theta)$ قابلة للاشتقاق الجزئي بالنسبة إلى كلٍّ من المْركَّبات $\theta_1$، $\theta_2$، ...و $\theta_k$ حيث يكون للدالَّة $L\,(\vec{x}_n;\theta)$ قيمة عظمى عندما تكون:

$$\frac{\partial L(\vec{x}_n\,;\,\theta)}{\partial\theta_i}=0 \qquad ;\forall\, i\in\mathbb{N}_k \qquad [10,5]$$

وحينئذ تُعيَّن مُقدِّرات الأرجحية العظمى $\hat{\theta}_1$، $\hat{\theta}_2$، ...و $\hat{\theta}_k$ الموافقة للمَعَالم $\theta_1$، $\theta_2$، ...و $\theta_k$ (على الترتيب) من خلال حل جملة المعادلات الآتية:

$$\frac{\partial L(\vec{x}_n\,;\,\theta)}{\partial\theta_1}\Bigg|_{\theta=\hat{\theta}}=0\ \ \&\ \ \frac{\partial L(\vec{x}_n\,;\,\theta)}{\partial\theta_2}\Bigg|_{\theta=\hat{\theta}}=0\ \ \&\ ...\ \&\ \ \frac{\partial L(\vec{x}_n\,;\,\theta)}{\partial\theta_k}\Bigg|_{\theta=\hat{\theta}}=0$$

وبعد ذلك نضع $\hat{\theta}=\left(\hat{\theta}_1,\hat{\theta}_2,...,\hat{\theta}_k\right)$.

٣- نعلم أنَّه إذا كانت $g$ دالَّة حقيقية موجبة ومُعرَّفة على $\mathbb{R}$، وكان $\frac{\partial g(x)}{\partial x}\Big|_{x=x_0}=0$، فإنَّه العلاقة $\frac{\partial \ln g(x)}{\partial x}\Big|_{x=x_0}=0$ ستكون محقَّقةً أيضاً، ومن ثمّ يُمكننا استخدام دالَّة اللوغاريتم الطبيعي لتعيين مُقدِّرات الأرجحية العظمى من خلال حل جملة المعادلات الآتية:

$$\frac{\partial \ln L\,(\vec{x}_n\,;\,\theta)}{\partial\theta}\Bigg|_{\theta=\hat{\theta}}=0 \qquad [10,6]$$

وفي حال كانت $\theta$ متَّجهاً $\theta=(\theta_1,\theta_2,...,\theta_k)$ فإنَّنا نستخدم العلاقات الآتية:

$$\left.\begin{array}{l}
\frac{\partial \ln L\,(\vec{x}_n\,;\,\theta)}{\partial\theta_1}\Big|_{\theta=\hat{\theta}}=0\\
\frac{\partial \ln L\,(\vec{x}_n\,;\,\theta)}{\partial\theta_2}\Big|_{\theta=\hat{\theta}}=0\\
\ \vdots \qquad \vdots \quad \vdots \qquad \vdots\\
\frac{\partial \ln L\,(\vec{x}_n\,;\,\theta)}{\partial\theta_k}\Big|_{\theta=\hat{\theta}}=0
\end{array}\right\} \qquad [10,7]$$

**(٥, ٢, ١, ١٠) أمثلة**

١- ليكن $\left[\Omega, \mathbb{S}, \mathcal{P}=\{P_\theta ; \theta \in \Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ دالَّة كثافته الاحتمالية مُعطاة من خلال العلاقة الآتية:

$$f_X(x\ ;\theta)=\frac{1}{\theta}\mathbf{e}^{-\frac{x}{\theta}}\ \mathbf{I}_{(0\ ,\ +\infty)}(x) \quad ; \forall x \in \mathbb{R}$$

ولتكن $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، ولنقم بتعيين دالَّة الأرجحية لهذه العيِّنة.

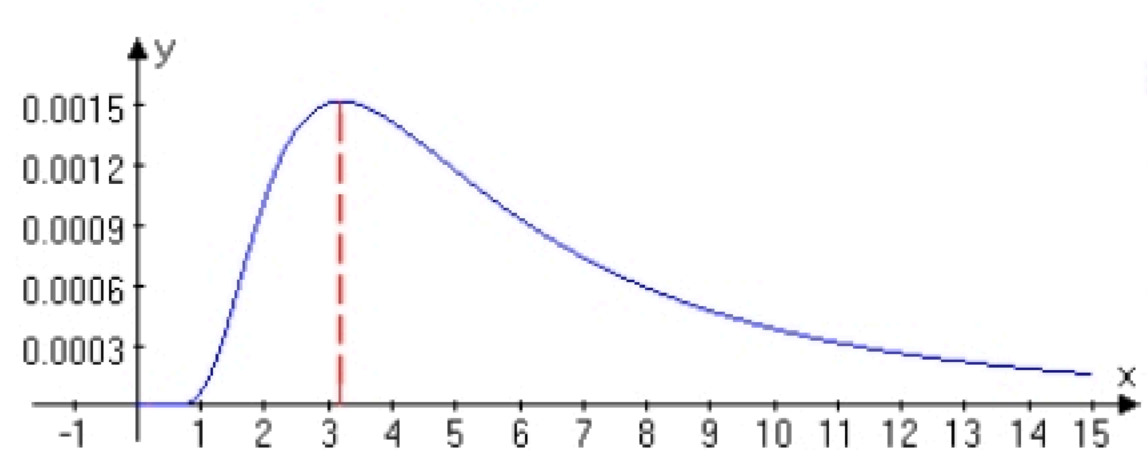
**الحل**: لدينا من تعريف دالَّة الأرجحية لعيِّنة ما يلي:

$$L(\vec{x}_n;\theta)=f_{\mathrm{X}_n}(\vec{x}_n;\theta)=\prod_{k=1}^{n}f_X(x_k;\theta)=\left(\frac{1}{\theta}\right)^n \mathbf{exp}\left[\frac{-1}{\theta}\sum_{k=1}^{n}x_k\right]\cdot\prod_{k=1}^{n}\mathbf{I}_{(0,\infty)}(x_k)$$

فعلى سبيل المثال لو أخذنا $x=3$ وبفرض أنَّ $\vec{x}_3=(2, 4, 3.6)$، فإنَّه سيكون لدينا:

$$L\left((2, 4, 3.6)\ ;\theta\right)=\frac{1}{\theta^3}\mathbf{e}^{-\frac{9.6}{\theta}}$$

حيث يمكن تبيان أنَّ لهذه الدالَّة قيمة عظمى عند النقطة $x=3.2$ (انظر الشكل الآتي).




الشكل (١, ١٠)

٢- ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً له توزيع منتظم مستمر على فترة $(a\ ,b) \subseteq \mathbb{R}$، ولتكن $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، ولنقم بتعيين مُقدِّرات الأرجحية العظمى لكلٍّ من $a$ و $b$.

**الحل**: من أجل ذلك لدينا:

$$L\left(\vec{x}_n;(a\ ,b)\right)=f_{\mathrm{X}_n}(\vec{x}_n;\theta)=\prod_{k=1}^{n}f_X(x_k;\theta)=\left(\frac{1}{b-a}\right)^n\cdot\prod_{k=1}^{n}\ \mathbf{I}_{(a\ ,b)}(x_k)$$

ومن ثمّ يكون لهذه الدالة قيمة عظمى عندما يكون $a \le x_1, x_2, ..., x_n \le b$، فإذا افترضنا أنَّ $\hat{a}$ و $\hat{b}$ هما مُقدِّرا الأرجحية العظمى لـ $a$ و $b$ على الترتيب، فإنَّه سيكون لدينا المتباينتين $\hat{a} \le x_1, x_2, ..., x_n$ و $\hat{b} \ge x_1, x_2, ..., x_n$ مُحقَّقتين، ومن ثمَّ لو قمنا بترتيب القيم الملحوظة لهذه العيِّنة من الأصغر إلى الأكبر على النحو $x_1^* > x_2^* > ... > x_n^*$، فعندئذ سيكون $\hat{a} \le x_1^* < x_2^* < ... < x_n^* \le \hat{b}$، وبالتالي ينتج لدينا أنَّ:

$$\hat{a}=X_1^*=\mathbf{min}\left\{X_1, X_2\ , ...\ , X_n\right\}$$

$$\hat{b}=X_n^*=\mathbf{max}\left\{X_1, X_2\ , ...\ , X_n\right\}$$

وبهذا يكون قد تمَّ تعيين مُقدِّرات مَعَالِم هذا المجتمع.

٣- ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع لتوزيع بواسون بمَعْلمة $\lambda > 0$، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، ولنقم بتعيين مُقدِّر الأرجحية العظمى للمَعْلمة $\lambda$.

الحـل: من أجل ذلك لدينا:

$$L(\vec{x}_n ;\lambda) = \prod_{k=1}^{n} f_X(x_k ;\lambda) = \mathbf{e}^{-n\cdot\lambda} \ \lambda^{\sum_{k=1}^{n} x_k} \left(\prod_{k=1}^{n} (x_k !)\right)^{-1} \qquad (*)$$

وبأخذ اللوغاريتم الطبيعي لطرفي العلاقة الأخيرة يصبح لدينا:

$$\ln L(\vec{x}_n ;\lambda) = -n\cdot\lambda + \left(\sum_{k=1}^{n} x_k\right)\ln\lambda - \sum_{k=1}^{n}\sum_{\ell=1}^{k} \ln x_\ell$$

حيث يجب أن يكون $\left.\dfrac{\partial \ln L(\vec{x}_n ;\lambda)}{\partial\lambda}\right|_{\lambda=\hat{\lambda}} = 0$، وباشتقاق العلاقة (*) بالنسبة إلى المَعْلمة $\lambda$ ينتج لدينا:

$$\left.\left(-n + \frac{1}{\lambda}\sum_{k=1}^{n} x_k\right)\right|_{\lambda=\hat{\lambda}} = 0 \ \Rightarrow\ \hat{\lambda} = \frac{1}{n}\sum_{k=1}^{n} x_k \ \Rightarrow\ \hat{\lambda} = \bar{X}$$

ومن ثمَّ يكون $\hat{\lambda} = \bar{X}$ هو مُقدِّر الأرجحية العظمى لمَعْلمة المجتمع البواسوني.

٤- ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع للتوزيع الطبيعي بمَعْلمتين $\mu \in \mathbb{R}$ و $\sigma > 0$، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، ولنقم بتعيين مُقدِّر الأرجحية العظمى لكلٍّ من $\mu$ و $\sigma$.

الحـل: من أجل ذلك لدينا:

$$L(\vec{x}_n ;(\mu,\sigma)) = \prod_{k=1}^{n} f_X(x_k ;(\mu,\sigma)) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \mathbf{exp}\left[-\frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-\mu)^2\right]$$

وبأخذ اللوغاريتم الطبيعي لطرفي العلاقة الأخيرة يصبح لدينا:

$$\ln L(\vec{x}_n ;(\mu,\sigma)) = -n\cdot\ln\sigma - n\cdot\ln\sqrt{2\pi} - \frac{1}{2\sigma^2}\sum_{k=1}^{n}(x_k-\mu)^2 \qquad (1)$$

حيث يجب أن يكون لدينا:

$$\left.\frac{\partial \ln L(\vec{x}_n ;(\mu,\sigma))}{\partial\mu}\right|_{(\mu,\sigma)=(\hat{\mu},\hat{\sigma})} = 0 \quad \& \quad \left.\frac{\partial \ln L(\vec{x}_n ;(\mu,\sigma))}{\partial\sigma}\right|_{(\mu,\sigma)=(\hat{\mu},\hat{\sigma})} = 0$$

وبالاشتقاق الجزئي لطرفي العلاقة (1) بالنسبة إلى المَعْلمة $\mu$ و $\sigma$ كلٌّ على حده يكون لدينا العلاقتين الآتيتين على الترتيب :

$$\left.\sum_{k=1}^{n}(x_k-\mu)\right|_{\mu=\hat{\mu}} = 0 \qquad \& \qquad \left.\sum_{k=1}^{n}\left(\frac{(x_k-\mu)^2}{\sigma^3} - \frac{n}{\sigma}\right)\right|_{\mu=\hat{\mu},\ \sigma=\hat{\sigma}} = 0 \qquad (2)$$

وبحل جملة المعادلتين (2) حلاً مشتركاً نحصل على النتائج الآتية:

$$\hat{\mu} = \frac{1}{n}\sum_{k=1}^{n} x_k \ \Rightarrow\ \hat{\mu} = \bar{X} \qquad \& \qquad \hat{\sigma}^2 = \frac{1}{n}\sum_{k=1}^{n}(x_k-\bar{x})^2 \ \Rightarrow\ \hat{\sigma}^2 = \tilde{S}^2 = \frac{n-1}{n}S^2$$

أي إنَّه لدينا مُقدِّر الأرجحية العظمى لـ $\mu$ هو $\hat{\mu} = \bar{X}$، وأمَّا مُقدِّر الأرجحية العظمى لـ $\sigma$ فهو $\hat{\sigma} = \tilde{S} = \sqrt{\dfrac{n-1}{n}}\, S$.

**(١٠,١,٢,٦) ملاحظات**

١- فيما يلي وعندما نذكر كلمة مقدِّر لمَعْلمة فإنَّه علينا الأخذ بالحسبان أنَّ هذا المقدِّر هو إحصاءة حتى لو لم يُذكر بهذا المسمَّى.

٢- فيما يلي سنورد بعض الملاحظات دون الخوض في تعليلاتها، حيث يمكن لمن يودّ الاطلاع على التعليلات الرجوع إلى المراجع ذات الصلة.

أ- إذا كان $\hat{\theta}$ هو مُقدِّر الأرجحية العظمى للمَعْلمة $\theta$ وكان لدينا $h = g(\theta)$ دالَّة وحيدة القيمة في $\theta$، فعندئذ يكون $\hat{h} = g(\hat{\theta})$ هو مُقدِّر الأرجحية العظمى للدالَّة $h$ أيضاً.

ب- في حال عدم تطابق نتائج طريقة العزوم مع نتائج طريقة الأرجحية العظمى في تعيين مُقدِّرات مَعَالم مجتمع ما، فإنَّنا نأخذ المقدِّرات التي نحصل عليها من طريقة الأرجحية العظمى؛ وذلك لأنَّه ثبت تجريبياً أنَّ مُقدِّرات الأرجحية العظمى هي أفضل من المقدِّرات التي نحصل عليها من طريقة العزوم.

ج- عندما يكون حجم العيِّنة كبيراً بقدر كاف فإنَّ سيكون لـ $\hat{\theta}$ (مُقدِّر الأرجحية العظمى للمَعْلمة $\theta$) توزيع طبيعي على وجه التقريب.

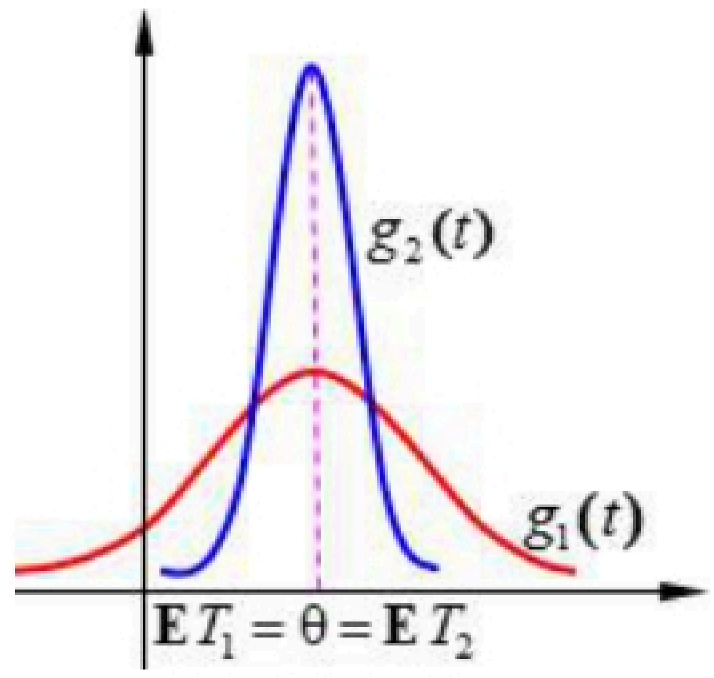
## (١٠,٢) التقدير النقطي

### Point Estimation

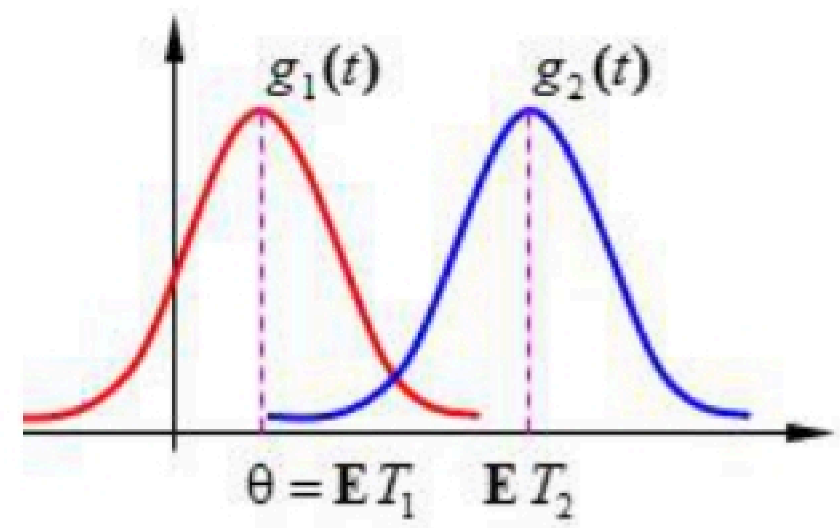
لقد قدَّمنا في الفقرة السابقة بعض الطرائق المستخدمة في تعيين مُقدِّر مَعْلمة مجتمع، فلو أمعنا النظر في مقدِّرات المَعَالم لوجدنا أنَّه من أجل عيِّنة مُحدَّدة $\mathfrak{X}_n$ ستكون قيم هذه المقدِّرات هي قيم تقديرية لمَعَالم المجتمع، وبدورها يمكن أن تُمثَّل بنقاط على المحور الحقيقي؛ ولذلك يُطلق على هذه المقدِّرات اسم **مُقدِّرات نقطية** لمَعَالم المجتمع، أي إنَّ هذه الطريقة في التقدير تهدف إلى تقدير قيمة المَعْلمة للمجتمع بقيمة مُحدَّدة تماماً. لكن قد يَحدث (في كثير من الحالات) أن يكون لدينا أكثر من مُقدِّر لمَعْلمة المجتمع، ومن ثمّ السؤال الذي يُطرح هنا: **أيٌّ من هذه المقدِّرات سنختار أو نفضل؟**

إنَّ السؤال السابق يعنى أنَّه أيٌّ من هذه المقدِّرات لمَعْلمة مجتمع يُعطينا استنباطاً أفضل للمعلومات حول هذه المَعْلمة؟

فلو افترضنا أنَّه لدينا الإحصاءة $T_1 = g_1(\mathfrak{X}_n)$ و $T_2 = g_2(\mathfrak{X}_n)$ بدالتي كثافة احتمالية $g_1(t)$ و $g_2(t)$ على الترتيب، فإنَّه من الممكن أن يكون توزيع أحدهما متمركزاً عند قيمة المَعْلمة $\theta$ في حين أنَّ الآخر متمركز بعيداً عن $\theta$ (انظر الشكل (١٠,٢.أ))، وحتى في حال تمركز توزيع كل منها عند القيمة $\theta$، فمن الممكن أن يكون تباين توزيع الإحصاءة $T_1$ أكبر من تباين توزيع الإحصاءة $T_2$، ومن ثمّ ستكون دالَّة الكثافة $g_1(t)$ منبسطة بالمقارنة مع دالَّة الكثافة $g_2(t)$ (انظر الشكل (١٠,٢.ب)).




الشكل (١٠,٢.ب)




الشكل (١٠,٢.أ)

إنَّ الإجابة على السؤال السابق تفرض علينا إعطاء معايير نُميِّز من خلالها صفات كل مُقدِّر ومن ثمَّ اختيار المقدِّر الأنسب لدراستنا.

### (١٠,٢,١) المقدِّرات غير المنحازة (أو المُنْصفة)

إنَّ أولى الصفات التي سنذكرها هيَ صفة عدم التحيز (أو الإنصاف) Unbiasedness التي سنُعرِّفها بعد تقديم ما يُعرف باسم مقْدار التَحيُّز لمُقدِّر مَعْلمة مجتمع.

#### (١٠,٢,١,١) تعريف (مقدار التحيُّز لمُقدِّر مَعْلمة مجتمع إحصائي) Unbiasedness Quantity of a Parameter

ليكن لدينا $[\Omega, \mathbb{S}, \mathcal{P}=\{P_\theta;\theta\in\Theta\}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولنأخذ $\mathcal{X}_n$ عيِّنة من هذا المجتمع، ولنفترض أنَّ $T = g(\mathcal{X}_n)$ مقدِّر للمَعْلمة $\theta$، فعندئذ يُعرَّف مقْدار تحيُّز المُقدِّر $T$ للمَعْلمة $\theta$ (وسنرمز له بـ $\vartheta(T,\theta)$) من خلال العلاقة الآتية:

$$\vartheta(T,\theta) := \mathbf{E}T - \theta \qquad [10,8]$$

#### (١٠,٢,١,٢) تعريف (المقدِّر غير المنحاز Unbiased Estimator)

ليكن لدينا $[\Omega, \mathbb{S}, \mathcal{P}=\{P_\theta;\theta\in\Theta\}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولنأخذ $\mathcal{X}_n$ عيِّنة من هذا المجتمع، فعندئذ يُقال عن إحصاءة $T = g(\mathcal{X}_n)$ إنَّها مُقدِّر غير منحاز للمَعْلمة $\theta$ إذا كانت العلاقة الآتية مُحقَّقة:

$$\vartheta(T,\theta) = 0 \qquad [10,9]$$

نلاحظ أنَّ العلاقة السابقة [10,9] تكافئ العلاقة $\mathbf{E}T = \theta$ التي تستخدم غالباً في حل الأمثلة والتمارين.

#### (١٠,٢,١,٣) أمثلة

١- ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع للتوزيع الطبيعي بمتوسط $\mu$ وانحراف معياري $\sigma$، ولنأخذ $\mathcal{X}_n = (X_1, X_2, ..., X_n)$ عيِّنة من هذا المجتمع، فعندئذ:

أ- نجد أنَّ الإحصاءة $\bar{X} = g(\mathcal{X}_n)$ (مقدِّرمتوسط المجتمع الطبيعي كما بيّنا ذلك سابقاً) هي مُقدِّر غير منحاز لمتوسط هذا المجتمع، وذلك لأنَّه بحسب العلاقة [10,9] لدينا $\mathbf{E}\bar{X} = \mu$.

ب- بفرض أنَّ:

$$\bar{X}_w = g(\mathcal{X}_n) = \alpha_1 X_1 + \alpha_2 X_2 + ... + \alpha_n X_n$$

علماً أنَّ $\alpha_1$، $\alpha_2$، ...و $0 < \alpha_n$ مع $\sum_{k=1}^{n} \alpha_k = 1$ (المتوسط الموزونللعيِّنة) هو مقدِّر لمتوسط هذا المجتمع، فعندئذ نجد أنَّ هذه الإحصاءة $\bar{X}_w$ هي مُقدِّر غير منحاز لمتوسط هذا المجتمع أيضاً، وذلك لأنَّه بسبب الاستقلال بين المتغيِّرات العشوائية $X_1$، $X_2$، ...و $X_n$ وتطابقها في التوزيع مع المتغيِّر العشوائي $X$ سيكون لدينا ما يلي مُحقَّقاً:

$$\mathbf{E}\bar{X}_w = \mathbf{E}\left(\sum_{k=1}^{n} \alpha_k X_k\right) = \sum_{k=1}^{n} \alpha_k \cdot \mathbf{E}X_k = \sum_{k=1}^{n} \alpha_k \cdot \mathbf{E}X = (\mathbf{E}X)\cdot \sum_{k=1}^{n} \alpha_k = \mathbf{E}X = \mu$$

ج- إنَّ الإحصاءة $\tilde{S}^2$ هي مُقدِّر منحاز لتباين هذا المجتمع $\sigma^2$، وذلك لأنَّه لدينا من العلاقة [9–12-a] ما يلي:

$$\mathbf{E}\tilde{S}^2 = \frac{n-1}{n}\sigma^2 \neq \sigma^2$$

ومقدار تحيُّزه $\vartheta(\tilde{S}^2, \sigma^2)$ يساوي:

$$\vartheta\left(\tilde{S}^2,\ \sigma^2\right)=\mathbf{E}\tilde{S}^2-\sigma^2=-\frac{\sigma^2}{n}$$

د- إنَّ الإحصاءة $S^2$ هي مُقدِّر غير منحاز لتباين هذا المجتمع $\sigma^2$، وذلك لأنَّ:

$$\mathbf{E}S^2=\frac{n}{n-1}\ \mathbf{E}\tilde{S}^2=\frac{n}{n-1}\frac{n-1}{n}\sigma^2=\sigma^2$$

٢- ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع للتوزيع المنتظم المستمر على فترة $[a,b]$ مع $-\infty<a<b<\infty$، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، ولننظر إن كان $\bar{X}$ و $\tilde{S}^2$ غير منحازين لـ $\mu$ متوسط و $\sigma^2$ تباين هذا المجتمع على الترتيب أم لا؟

في الواقع لدينا من استقلال المتغيِّرات العشوائية $X_1$، $X_2$، ...و $X_n$، وبسبب تطابقها في التوزيع مع المتغيِّر العشوائي $X$ ما يلي مُحقَّقاً:

$$\mathbf{E}\bar{X}=\frac{1}{n}\mathbf{E}\sum_{k=1}^{n}X_k=\frac{1}{n}n\cdot\mathbf{E}X=\mathbf{E}X=\frac{a+b}{2}=\mu$$

وهذا يعني أنَّ $\bar{X}$ هو مُقدِّر غير منحاز لمتوسط هذا المجتمع. كذلك للأسباب التي ذُكرت سابقاً نفسها نجد ما يلي:

$$\mathbf{E}\tilde{S}^2=\frac{1}{n}\mathbf{E}\sum_{k=1}^{n}\left(X_k-\bar{X}\right)^2=\mathbf{E}X^2-\mathbf{E}\bar{X}^2$$

$$=\frac{b^2+ab+a^2}{3}-\frac{a^2+2ab+b^2}{4}=\frac{b^2-2ab+a^2}{12}=\frac{(b-a)^2}{12}=\mathbf{var}\,X=\sigma^2$$

وهكذا نجد أنَّ الإحصاءة $\tilde{S}^2$ هي مُقدِّر غير منحاز لتباين هذا المجتمع في حين أنَّنا نجد أنَّ الإحصاءة $S^2$ هو مُقدِّر منحاز لتباين هذا المجتمع؛ وذلك لأنَّ:

$$\mathbf{E}S^2=\frac{n}{n-1}\mathbf{E}\tilde{S}^2=\frac{n}{n-1}\sigma^2\neq\sigma^2$$

### (٤ ,١ ,٢ ,١٠) نتيجة

نستنتج من المثالين السابقين أنَّ الإحصاءة التي تكون منصفة لمجتمع معيَّن قد لا تكون كذلك بالنسبة إلى مجتمع آخر.

الآن لو أمعنا النظر في الفقرة (ج) من المثال /١/ في (٣ ,١ ,٢ ,١٠)، فإنَّنا نلاحظ أنَّه عندما يسعى حجم العيِّنة $n$ إلى اللانهاية فإنَّ الإحصاءة $\tilde{S}^2$ تسعى لتصبح مُقدِّراً غير منحاز لـ $\sigma^2$. إنَّ الإحصاءات التي تتمتَّع بهذه الصفة يُقال عنها إنَّها تتمتَّع بالصفة **التقاربية** Asymptotic Property، والتعريف الآتي يقدِّم لنا هذا النوع من المقدِّرات.

### (٥ ,١ ,٢ ,١٠) تعريف (المقدِّر غير المنحاز تقاربيًّا Asymptotic Unbiased Estimator)

ليكن لدينا $\left[\Omega, \mathbb{S}, \mathcal{P}=\{P_\theta;\theta\in\Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، و $T=g(\mathfrak{X}_n)$ مُقدِّراً للمَعْلمة $\theta$، فإذا كان:

$$\lim_{n\to\infty}\mathbf{E}g\left(\mathfrak{X}_n\right)=\theta \qquad [10,10]$$

فعندئذ يُقال عن الإحصاءة $T$ إنَّها مُقدِّر **غير منحاز تقاربيًّا** للمَعْلمة $\theta$.

### (٦ ,١ ,٢ ,١٠) أمثلة

١- بالعودة إلى المثال /١/ من (٣ ,١ ,٢ ,١٠) نجد أنَّ الإحصاءة $\tilde{S}^2$ هي مُقدِّر غير منحاز تقاربياً لتباين المجتمع الطبيعي

المُعطى، وذلك لأنَّ:

$$\lim_{n\to\infty} \mathbf{E}\tilde{S}^2 = \lim_{n\to\infty} \frac{n-1}{n}\mathbf{E}S^2 = \lim_{n\to\infty} \frac{n-1}{n}\sigma^2 = \sigma^2 \lim_{n\to\infty} \frac{n-1}{n} = \sigma^2$$

٢- بالعودة إلى المثال / ٢/ من (١٠,٢,١,٣) نجد أنَّ الإحصاءة $S^2$ هي مُقدِّر غير منحاز تقاربياً لتباين المجتمع المنتظم المستمر على الفترة $(a , b)$، وذلك لأنَّ:

$$\lim_{n\to\infty} \mathbf{E}S^2 = \lim_{n\to\infty} \frac{n}{n-1}\mathbf{E}\tilde{S}^2 = \lim_{n\to\infty} \frac{n\,(b-a)^2}{12(n-1)} = \lim_{n\to\infty} \frac{n}{n-1}\sigma^2 = \sigma^2 \lim_{n\to\infty} \frac{n}{n-1} = \sigma^2$$

### (١٠,٢,٢) المقدِّرات المتَّسقة Consistent Estimators

لقد لوحظ أنَّ تباين بعض المقدِّرات يصغر بكبر حجم العيِّنة بحيث إنَّه عندما يَسعى حجم العيِّنة إلى اللانهاية فإنَّ تباين العيّنة سوف يسعى إلى الصفر. إنَّ هذه الخاصية تُعْرف باسم **شرط ماركوف** الذي تقدِّمه لنا الفقرة الآتية.

### (١٠,٢,٢,١) شرط ماركوف Markov Condition

ليكن لدينا $[\Omega ,\mathbb{S},\mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولنأخذ $\mathcal{X}_n$ عيِّنة من هذا المجتمع، و $T = g\,(\mathcal{X}_n)$ إحصاءة مُعطاة. عندئذ يُقال إنَّ $T$ تُحقِّق شرط ماركوف **إذا وفقط إذا** تحقَّقت العلاقة الآتية:

$$\lim_{n\to\infty} \mathbf{var}\, g\,(\mathcal{X}_n) = 0 \qquad [10,11]$$

### (١٠,٢,٢,٢) مثال

ليكن لدينا $[\Omega ,\mathbb{S},\mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ له توزيع ما، ويملك عزوماً ابتدائية حتى المرتبة الرابعة على الأقل، وتباينه $\sigma^2$، ولتكن $\mathcal{X}_n$ عيِّنة من هذا المجتمع، ولنأخذ الإحصاءة $\tilde{S}^2$. عندئذ نجد أنَّ $\tilde{S}^2$ يُحقِّق شرط ماركوف، وذلك لأنَّه من العلاقة [9–13-a] لدينا:

$$\lim_{n\to\infty} \mathbf{var}\, \tilde{S}^2 = \lim_{n\to\infty} \left( \frac{\mu^4-\sigma^4}{n} - \frac{2(\mu^4-2\sigma^4)}{n^2} + \frac{\mu^4-3\sigma^4}{n^3} \right) = 0$$

توجد في الواقع مقدِّرات تتمتَّع بصفة عدم التّحيُّز تقاربياً وتحقِّق شرط ماركوف أيضاً. إنَّ هذا النوع من المقدِّرات يُقال عنها إنَّها تتمتَّع بـ خاصيّة الاتّساق Consistency Property، والمقدَّر الذي يجمع بين هاتين الصفتين يُدعى **مُقدِّراً متَّسقاً**، والفقرة الآتية تقدِّم لنا تعريفه.

### (١٠,٢,٢,٣) تعريف (المقدِّر المتَّسق Consistent Estimator)

ليكن لدينا $\left[\Omega ,\mathbb{S},\mathcal{P}=\{P_\theta ;\theta\in\Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولتكن $\mathcal{X}_n$ عيِّنة من هذا المجتمع، فإذا كان $T = g\,(\mathcal{X}_n)$ مُقدِّراً للمَعْلمة $\theta$، فعندئذ يُقال عن الإحصاءة $T$ إنَّها **مُقدِّر متَّسق** للمَعْلمة $\theta$ **إذا وفقط إذا** تحقَّقت العلاقتين:

$$\textbf{a)}\ \lim_{n\to\infty} \mathbf{E}\big(g\,(\mathcal{X}_n)\big) = \theta \qquad \& \qquad \textbf{b)}\ \lim_{n\to\infty} \mathbf{var}\big(g\,(\mathcal{X}_n)\big) = 0$$

### (١٠,٢,٢,٤) مثال

ليكن لدينا $[\Omega ,\mathbb{S},\mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ له توزيع ما، ويملك عزوماً ابتدائية حتى المرتبة الرابعة على الأقل، وتباينه $\sigma^2$، ولتكن $\mathcal{X}_n$ عيِّنة من هذا المجتمع، فنجد من المثالين (١٠,٢,١,٦) و(١٠,٢,٢,٢) أنَّ الإحصاءة $\tilde{S}^2$ هي مُقدِّر غير منحاز تقاربيًّا لتباين هذا المجتمع، ويُحقِّق شرط ماركوف أيضاً، ومن ثمَّ الإحصاءة $\tilde{S}^2$ هي مُقدِّر متَّسق للمَعْلمة $\sigma^2$.

### (١٠,٢,٢,٥) ملاحظات

١ – تجدر الملاحظة إلى أنّه توجد نماذج أخرى من المقدِّرات المتّسقة التي تُعرف بالمقدِّرات الضعيفة الاتّساق والمُقدِّرات القوية الاتّساق، ولكن لن نخوض في هذه النماذج من المقدرات.

٢ – إذا كان حجم العيِّنة كبيراً، فإنَّ مُقدِّرات الأرجحية العظمى لمَعْلمة التباين $\sigma^2$ ستكون متّسقة أيضاً.

### (١٠,٢,٣) المقدِّرات الكفؤة Efficient Estimators

لقد لاحظنا سابقًا أنّه من الممكن أن يكون لدينا أكثر من مُقدِّر غير منحاز لمَعْلمة مجتمع مُعطى، والسؤال الذي يُطرح هنا:

**أيُّ المقدِّرات غير المنحازة سنفضِّل؟**

الفقرات الآتية تقدِّم مفاهيم تساعدنا في اختيار المقدِّر غير المنحاز الذي يُعطينا معلومات أفضل حول مَعْلمة المجتمع.

#### (١٠,٢,٣,١) تعريف (المقدِّر الكفؤ Efficient Estimator)

ليكن $\left[\Omega,\mathbb{S},\mathcal{P}=\{P_\theta;\theta\in\Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ بدالّة كثافة $f_X(x;\theta)$ قابلة للاشتقاق الجزئي بالنسبة إلى $\theta$، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، وبفرض أنَّ $T=g(\mathfrak{X}_n)$ مُقدِّر غير منحاز للمَعْلمة $\theta$، فعندئذ يُقال عن الإحصاءة $T$ إنَّها مُقدِّر **كفؤ** للمَعْلمة $\theta$ **إذا وفقط إذا** تحقَّقت العلاقة الآتية:

$$\mathbf{var}T=\left[n\cdot\mathbf{E}\left(\frac{\partial\,\mathbf{ln}f_X(x\;;\theta)}{\partial\theta}\right)^2\right]^{-1} \qquad [10,12]$$

وأمّا النسبة المكوِّنة للمقدار الآتي:

$$\mathbf{ef}\,(T):=\frac{1}{\mathbf{var}T}\left[n\cdot\mathbf{E}\left(\frac{\partial\,\mathbf{ln}f_X(x\;;\theta)}{\partial\theta}\right)^2\right]^{-1} \qquad [10,13]$$

فإنَّها تُدعى **مقدار كفاءة** المقدِّر $T$.

#### (١٠,٢,٣,٢) تعريف (المقدِّر الأكفأ من مُقدِّر آخر Estimator more efficient from another)

ليكن $\left[\Omega,\mathbb{S},\mathcal{P}=\{P_\theta;\theta\in\Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولتكن $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، وليكن $T_1=g_1(\mathfrak{X}_n)$ و $T_2=g_2(\mathfrak{X}_n)$ مقدِّرين غير منحازين للمَعْلمة $\theta$، فإذا كان $\mathbf{var}T_2\ge\mathbf{var}T_1$، فعندئذ يُقال إنَّ المقدِّر $T_1$ **أكفأ** من المقدِّر $T_2$.

تُدعى النسبة $\dfrac{\mathbf{var}T_1}{\mathbf{var}T_2}$ بـ **الكفاءة النسبية** Relative Efficiency لـ $T_2$ بالنسبة إلى $T_1$، وسنرمز لها بـ $\mathbf{ef}\,(T_1,T_2)$، أي إنَّ:

$$\mathbf{ef}\,(T_1,T_2):=\frac{\mathbf{var}T_1}{\mathbf{var}T_2} \qquad [10,14]$$

#### (١٠,٢,٣,٣) تعريف (المقدِّر الأعلى كفاءة Highest Efficient Estimator):

ليكن $\left[\Omega,\mathbb{S},\mathcal{P}=\{P_\theta;\theta\in\Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، ولنفترض أنَّ $\mathfrak{T}$ هي أسرة كل المقدِّرات غير المنحازة للمَعْلمة $\theta$، فإذا وجد مُقدِّر $T\in\mathfrak{T}$ بحيث تكون المتباينة الآتية محقَّقة:

$$\mathbf{var}T\le\mathbf{var}T^{*}\quad;\forall\,T^{*}\in\mathfrak{T} \qquad [10,15]$$

فعندئذ يُقال عن $T$ إنَّه **المقدِّر الأعلى كفاءةً** للمَعْلمة $\theta$ (أو الأكفأ أو أفضل مُقدِّر خطي للمَعْلمة $\theta$).

من المتباينات المهمّة في مجال دراسة المقدِّرات النقطية ما يُعرف باسم متباينة راو-كرامر التي تقدِّمها المبرهنة الآتية (سنقبلها دون برهان).

**(٤ , ٢ , ١٠) مبرهنة (متباينة راو- كرامر Rao-Cramer Inequality)**

ليكن $\left[\Omega, \mathbb{S}, \mathcal{P}=\{P_\theta;\theta\in\Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي نظاميّ $X$ بالنسبة إلى $\theta \ni \Theta_o \supseteq \Theta$، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، و $T = g(\mathfrak{X}_n)$ مُقدِّراً نظاميًّا بالنسبة إلى $\theta \ni \Theta_o$، فعندئذ إذا كانت $\Theta_o$ مجموعة مفتوحة، فإنَّه ستكون المتباينة الآتية مُحقَّقة:

$$\mathbf{var}T \geq \left[1+\left(\frac{d\,\vartheta(T\ ,\theta)}{d\,\theta}\right)^2\right]\cdot\left[n\cdot\mathbf{E}\left(\frac{\partial\,\mathbf{ln}f_X\,(x\ ;\theta)}{\partial\theta}\right)^2\right]^{-1} \qquad [10,16]$$

وهذه المتباينة تُعرف باسم **متباينة راو-كرامر** نسبة إلى الإحصائي البريطاني (هندي الأصل) راو Calyampudi Radhakrishna Rao (1920) والرياضياتي الإحصائي السويدي كرامر (1893-1985) Harald Cramér.

إنَّ الطرف الأيمن من المتباينة السابقة يُدعى **الحد الأدنى لمتباينة راو-كرامر**، وإذا حقَّقت الإحصاءة $T$ علاقة المساواة في هذه المتباينة، فعندئذ يُقال إنَّ الإحصاءة $T$ **بلغت** الحد الأدنى لمتباينة راو-كرامر.

**(١ , ٤ , ٢ , ١٠) نتائج**

١- إذا كان $T$ مقدِّراً غير منحاز للمَعْلمة $\theta$، فعندئذ بحسب تعريف المقدِّر المنصف سيكون لدينا $\vartheta(T\ ,\theta)=0$ ، ومن ثَمّ يصبح لمتباينة راو كرامر العرض الآتي:

$$\mathbf{var}T \geq \left[n\cdot\mathbf{E}\left(\frac{\partial\,\mathbf{ln}f_X\,(x\ ;\theta)}{\partial\theta}\right)^2\right]^{-1} \qquad [10,17]$$

ومنه ينتج أنَّ المقدِّر الكفؤ للمَعْلمة $\theta$ هو ذلك المقدِّر غير المنحاز الذي يُحقِّق علاقة المساواة في متباينة راو-كرامر (أي يَبلغ الحد الأدنى لمتباينة راو-كرامر).

٢- إذا كان $T$ مُقدِّراً كفؤًا للمَعْلمة $\theta$، فعندئذ سيكون لدينا ما يلي مُحقَّقاً:

$$\mathbf{var}T = \left[\mathbf{I}_{\mathfrak{X}_n}(\theta)\right]^{-1} \qquad [10,18]$$

علماً أنَّ $\mathbf{I}_{\mathfrak{X}_n}(\theta)$ هي كمية معلومات فيشر التي تقدمها العيِّنة $\mathfrak{X}_n$ حول المَعْلمة $\theta$.

**(٢ , ٤ , ٢ , ١٠) مثال**

١- ليكن لدينا $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع لتوزيع برنولي بمَعْلمة $0<p<1$، و $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، وبفرض أنَّ $T = \sum_{k=1}^{n}\alpha X_k$ إحصاءة مُعطاة، ولنقم بما يلي:

أ- حساب قيمة الثابت $\alpha$ بحيث يصبح الإحصاءة $T$ مقدِّراً غير منحاز للمَعْلمة $p$.

ب- تبيان إن كان $T$ مقدِّراً كُفؤًا للمَعْلمة $p$ أم لا.

ج- تبيان إن كان $T$ هو المقدِّر الأكفأ للمَعْلمة $p$ أم لا.

الحــل: من أجل الطلب:

أ – لكي يكون $T$ مقدِّراً غير منحاز للمَعْلمة $p$ يجب أنَّ يكون $\mathbf{E}T = p$، ومن ثمّ نجد الآتي:

$$p = \mathbf{E}T = \alpha \sum_{k=1}^{n} \mathbf{E}X_k = \alpha n \ \mathbf{E}X = \alpha n p$$

والتي ينتج عنها أنَّه حتى تصبح الإحصاءة $T$ مقدِّراً غير منحاز للمَعْلمة $p$ يجب أن تكون قيمة الثابت $\alpha$ تساوي $\frac{1}{n}$، ومن ثَمّ يجب أنَّ يكون $T = \overline{\pi}$.

ب– من أجل أي $x \in \{0,1\}$ لدينا $f_X(x\ ;\ p) = p^x (1-p)^{1-x}$، وبأخذ لوغاريتم الطرفين لهذه العلاقة ينتج لدينا:

$$\ln f_X(x\ ;\ p) = x \cdot \ln p + (1-x) \cdot \ln(1-p)$$

وبالاشتقاق الجزئي للعلاقة الأخيرة بالنسبة إلى $p$ نجد الآتي:

$$\frac{\partial \ln f_X(x\ ;\ p)}{\partial p} = \frac{x}{p} - \frac{1-x}{1-p}$$

وبالاشتقاق الجزئي للعلاقة الأخيرة بالنسبة إلى $p$ مرَّةً أخرى ينتج لدينا ما يلي:

$$\frac{\partial 2 \ln f_X(x\ ;p)}{\partial p^2} = \frac{-x}{p^2} - \frac{1-x}{(1-p)^2}$$

ومن ثَمّ يصبح لدينا:

$$\mathbf{E}\left(\frac{-\partial 2 f_X(x\ ;\ p)}{\partial p^2}\right) = \sum_{x=0}^{1} \frac{-\partial 2 f_X(x\ ;\ p)}{dp^2} f_X(x\ ;\ p) = \frac{1}{p(1-p)}$$

ومنه تكون كمية المعلومات التي تقدمها العيِّنة $\mathfrak{X}_n$ حول المَعْلمة $p$ هي:

$$\left[n \cdot \mathbf{E}\left(\frac{\partial \ln f_X(x\ ;\ p)}{\partial p}\right)^2\right]^{-1} = \frac{p(1-p)}{n} \qquad \mathbf{(1)}$$

وهكذا نجد أنَّ المقدِّر $\overline{\pi}$ يُحقِّق علاقة المساواة في متباينة راو–كرامر، ومن ثَمَّ $\overline{\pi}$ هو مقدِّر كُفؤ للمَعْلمة $p$.

ج– الآن لتبيان إن كان $\overline{\pi}$ هو المقدِّر الأكفأ للمَعْلمة $p$ أم لا سنقوم بحساب $\mathbf{var}T$، فنجد أنَّ:

$$\mathbf{var}T = \mathbf{var}\overline{X} = \frac{1}{n^2} \sum_{k=1}^{n} \mathbf{var}X_k = \frac{1}{n^2} n \cdot \mathbf{var}X = \frac{p(1-p)}{n} \qquad \mathbf{(2)}$$

والآن من العلاقتين **(1)** و **(2)** نستنتج أنَّ:

$$\mathbf{var}T = \left[n \cdot \mathbf{E}\left(\frac{\partial \ln f_X(x\ ;\ p)}{\partial p}\right)^2\right]^{-1}$$

وهذا يعني أنَّ الإحصاء $\overline{\pi}$ هو المقدِّر الأكفأ للمَعْلمة $p$ أيضاً.

تعطينا المبرهنة الآتية (التي سنقبلها دون برهان) الشرط اللازم والكافي حتى يَبلغ مُقدِّر مَعْلمة مجتمع إحصائي الحد الأدنى للمتباينة راو–كرامر.

**(١٠,٢,٤,٣) مبرهنة**

ليكن لدينا $\left[\Omega,\mathbb{S},\mathcal{P}=\{P_\theta;\theta\in\Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، و $T=g(\mathfrak{X}_n)$ مقدِّراً غير منحاز للمَعْلمة $\theta$ **بأقل تباين**، فعندئذ يبلغ $T$ الحد الأدنى لمتباينة راو-كرامر **إذا وفقط إذا** كانت الدالَّة $f_X(x;\theta)$ **أُسيَّة**، أي إنَّه يمكن كتابة الدالَّة $f_X(x;\theta)$ على النحو الآتي:

$$f_X(x;\theta)=\exp\left[s(\theta)+h(x)+u(\theta)\cdot v(x)\right] \qquad [10,17]$$

وتكون حينئذ الإحصاءة $U=\frac{1}{n}\sum_{k=1}^{n}\ln X_k$ هي مُقدِّر غير منحاز بأقل تباين للمَعْلمة $\frac{-s'(\theta)}{u'(\theta)}$، علماً أنَّ الاشتقاق هنا يتمُّ بالنسبة إلى المَعْلمة $\theta$.

**(١٠,٢,٤,٤) مثال**

ليكن $\left[\Omega,\mathbb{S},\mathcal{P}=\{P_\theta;\theta\in\Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ بدالَّة كثافة احتمالية $f_X(x;\theta)$ مُعطاة من خلال العلاقة:

$$f_X(x;\theta)=\theta\cdot x^{\theta-1}\mathbf{I}_{(0,1)}(x) \qquad ;\forall x\in\mathbb{R}$$

ولنقم بإيجاد مُقدِّر غير منحاز بأقل تباين للمَعْلمة $\frac{1}{\theta}$ مستخدمين في ذلك الحد الأدنى لمتباينة راو-كرامر.

من أجل ذلك نلاحظ أنَّه يمكننا كتابة الدالَّة $f_X(x;\theta)$ على النحو الآتي:

$$f_X(x;\theta)=\exp\left[\ln\theta-\ln x+\theta\ln x\right]\cdot\mathbf{I}_{(0,1)}(x)$$

وكما هو واضح فإنَّ لـ $f_X(x;\theta)$ نموذج دالَّة أُسيَّة، حيث لدينا:

$$s(\theta)=\ln\theta \quad \& \quad h(x)=-\ln x \quad \& \quad u(\theta)=\theta \quad \& \quad v(x)=\ln x$$

ومن ثمَّ بحسب المبرهنة السابقة تكون الإحصاءة الآتي:

$$T=\frac{-1}{n}\sum_{k=1}^{n}\ln X_k=\left(\ln\prod_{k=1}^{n}X_k\right)^{-\frac{1}{n}}$$

هي مُقدِّر غير منحاز بأقل تباين للمَعْلمة $\frac{-s'(\theta)}{u'(\theta)}=\frac{1}{\theta}$.

## (١٠,٣) فترات الثقة

### Confidence Intervals

لقد بيَّنا في دراستنا للمقدِّرات النقطية متى نفضل مقدِّر على آخر بخصوص مَعْلمة مجتمع معتمدين في ذلك على عيِّنة مسحوبة من المجتمع الذي قيد الدراسة. إنَّ عملية التفضيل لهذا المقدِّر على ذاك كانت مبنيَّة على بعض الخصائص التي يتمتَّع بها المقدِّر، ومن ثمَّ نكون على قدر ما من الثِّقة من أنَّ قيمة هذا المقدِّر قريبة من قيمة المَعْلمة المفترضة، ولكنَّ القيمة التي تنتج عن هذا المقدِّر كعدد مجرَّد لا تُعطينا القيمةَ الاحتمالية التي تدلُّ على مساواة قيمة هذا المقدِّر لقيمة المَعْلمة المقدَّرة. بمعنى آخر، ما هو احتمال أن تكون قيمة هذا المقدِّر الذي تمَّ تعيينه مساويةً أو قريبةً من قيمة هذه المَعْلمة المُقدَّرة؟

**(١٠,٣,١) مفهوم فترة الثِّقة لمَعْلمة مجتمع إحصائي Concept of Confidence Interval of Population**

من أجل توضيح ذلك سنأخذ $\left[\Omega,\mathbb{S},\mathcal{P}=\{P_\theta;\theta\in\Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولنسحب عيِّنة $\mathfrak{X}_n$ من هذا المجتمع، ولنفترض أنَّ $T=g(\mathfrak{X}_n)$ مُقدِّر للمَعْلمة $\theta$، فعندئذ الإجابة على السؤال السابق تكمن في تعيين قيمة

احتمال الحادث $\{\omega \in \Omega \,;\, |T(\omega)-\theta| < \varepsilon\}$ علماً أنَّ $\varepsilon > 0$ مُعطى. بكلمات أخرى، بفرض أنَّ $0 \le \alpha \le 1$ عدد حقيقي مُعطى، فما هي قيمة $\varepsilon > 0$ التي من أجلها تكون العلاقة الآتية محقَّقةً:

$$P(\,|T-\theta| < \varepsilon\,) = 1-\alpha \qquad [10,18\text{-a}]$$

وهنا نلاحظ أنَّه كلّما كانت قيمة $\alpha$ و $\varepsilon$ صغيرة، دلَّ ذلك على أنَّ قيمة المقدِّر $T$ ستكون قريبة من قيمة المَعْلمة المقدَّرة $\theta$ باحتمال كبير. إنَّ العلاقة الأخيرة تعطينا العلاقة الآتية:

$$P\left(T-\varepsilon < \theta < T+\varepsilon\right) = 1-\alpha \qquad [10,18\text{-b}]$$

ومن ثمّ تؤول المسألة إلى إيجاد مُقدِّرين $T_1 = T-\varepsilon$ و $T_2 = T+\varepsilon$ بحيث يكون احتمال أنَّ تقع قيمة المَعْلمة $\theta$ في فترة مفتوحة $(t_1, t_2)$ يساوي $1-\alpha$، علماً أنَّ $t_1$ و $t_2$ هما قيمة $T_1$ و $T_2$ على الترتيب، وفي هذه الحالة تُدعى الإحصاءة $T_1$ **حد الثّقة الأدنى** Lower Confidence Limit للمَعْلمة $\theta$ في حين تُدعى الإحصاءة $T_2$ **حد الثّقة الأعلى** Upper Confidence Limit للمَعْلمة $\theta$ .

### (١, ١, ٣, ١٠) تعريف (فترة الثّقة لمَعْلمة مجتمع Confidence Interval of Population)

ليكن $\left[\Omega, \mathcal{S}, \mathcal{P}=\{P_\theta ; \theta \in \Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ ، ولتكن $\mathfrak{X}_n$ عيِّنة مسحوبة من هذا المجتمع، وليكن $T_1 = g_1(\mathfrak{X}_n)$ و $T_2 = g_2(\mathfrak{X}_n)$ إحصائيتين مُعطيين، و $0 \le \alpha \le 1$، فعندئذ إذا كان $X$ مستمرًّا فإنَّ العلاقة الآتية:

$$P\left(T_1 < \theta < T_2\right) = 1-\alpha \qquad [10,19]$$

تُعرِّف لنا **$100\,(1\text{-}\alpha)\%$ فترة ثقة** للمَعْلمة $\theta$، وحينئذ المقدار $(1-\alpha)$ يُدعى **مستوى الثّقة** (أو مُعَامل الثّقة) Confidence Level.

### (٢, ١, ٣, ١٠) ملاحظات

١- في دراستنا المقبلة سيتم التركيز على المجتمعات الطبيعية والبرنولية، أي إنَّنا سنأخذ $X$ متغيِّراً عشوائياً طبيعيًّا (أو طبيعيًّا على وجه التقريب) أو برنو ليًّا، ومن أجل المجتمعات البرنولية سنأخذ عيِّنات كبيرة بقدر كاف بحيث يمكن تطبيق مبرهنة النهاية المركزية لـ موافير-لابلاس.

٢- من العلاقة [10-19] ومن أجل $\alpha_1 + \alpha_2 = \alpha$ يمكننا أنَّ نكتب العلاقتين الآتيتين:

$$P\left(T_1 \ge \theta\right) = \alpha_1 \qquad \& \qquad P\left(T_2 \le \theta\right) = \alpha_2$$

وفي الحالة الخاصة عندما يكون $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$ فإنَّه يُبرهن على أنَّ طول فترة الثّقة لـ $\theta$ سيكون أصغرياً (أي نحصل على أقصر طول فترة ثقة لـ $\theta$ )، ولذلك سينصب اهتمامنا فيما يأتي على إعطاء فترات الثّقة للمَعْلمة $\theta$ (أو لإحدى مُركِّبات $\theta$ إذا كان $\theta$ متَّجهًا من $\mathbb{R}^k$ ) من أجل الحالة التي يكون فيها $\alpha_1 = \alpha_2 = \frac{\alpha}{2}$، وذلك دون الخوض في كيفية استنتاج هذه الفترات، حيث يمكن للقارئ أنَّ يستنتج بعضها بسهولة من خلال تطبيق العلاقة [10-18-b] ومن ثمّ حساب قيمة $\varepsilon$ .

٣- إذا استطعنا تعيين $100\,(1-\alpha)\%$ فترة ثقة للمَعْلمة $\theta$ وفقاً للعلاقة [10-18-b] فعندئذ يمكننا تعيين $100\,(1-\alpha)\%$ فترة ثقة لأيّة دالَّة مطَّردة $g(\theta)$ من خلال العلاقة الآتية:

$$P\left(g(T_1) < g(\theta) < g(T_1)\right) = 1-\alpha \qquad [10,20]$$

فيما يلي سنقوم بتعيين فترات الثّقة في حال أنَّه لدينا مجتمع وحيد، وبعد ذلك في حال لدينا مجتمعين مستقلَّين مع الأخذ بالملاحظات / ٢, ١, ٣, ١٠ .

### (١٠,٣,٢) فترات الثّقة لَمْعلمتي مجتمع طبيعي

سنفترض على مدى هذه الفقرة أنّه لدينا $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع للتوزيع الطبيعي بمتوسط $\mu$ وانحراف معياري $\sigma$، وأنَّ $\mathfrak{X}_n = (X_1, X_2, ..., X_n)$ عيِّنة من هذا المجتمع بمتوسط $\overline{X}$ وانحراف معياري $S$.
تحت هذه الافتراضات سنقوم بتعيين فترة الثّقة لـ $\mu$ في كل من الحالتين الخاصّتين الآتيتين:

أ- إذا كان الانحراف المعياري $\sigma$ للمجتمع معلوماً.

ب- إذا كان الانحراف المعياري $\sigma$ للمجتمع مجهولاً.

وبعد ذلك سننتقل إلى تعيين فترة الثّقة لـ $\sigma^2$ (تباين هذا المجتمع). إنَّ المبرهنة الآتية (التي سنقبلها دون برهان) تقدِّم لنا فترات الثّقة لمتوسط مجتمع طبيعي.

#### (١٠,٣,٢,١) مبرهنة

تحت الفرضيات المذكورة آنفًا يكون لدينا ما يلي:

١- إذا كان الانحراف المعياري للمجتمع معلوماً مع $\sigma = \sigma_o$، فعندئذ تُعطى $100\,(1-\alpha)\,\%$ فترة ثقة لـ $\mu$ (متوسط هذا المجتمع) من خلال الفترة الآتية:

$$\left(\overline{x} - \frac{\sigma_o}{\sqrt{n}}\,\mathfrak{z}_{1-\frac{\alpha}{2}}\ ,\ \overline{x} + \frac{\sigma_o}{\sqrt{n}}\,\mathfrak{z}_{1-\frac{\alpha}{2}}\right) \qquad [10,21\text{-a}]$$

علماً أنَّ $\mathfrak{z}_k$ هي تلك النقطة على المحور الحقيقي التي يقع على يسارها وتحت منحنى الكثافة للتوزيع الطبيعي المعياري مساحة مقدارها يساوي $k$.

٢- إذا كان الانحراف المعياري للمجتمع $\sigma$ مجهولاً، فعندئذ تُعطى $100\,(1-\alpha)\,\%$ فترة ثقة لـ $\mu$ من خلال الفترة الآتية:

$$\left(\overline{x} - \frac{s}{\sqrt{n}}\,t_{1-\frac{\alpha}{2}}^{[n-1]}\ ,\ \overline{x} + \frac{s}{\sqrt{n}}\,t_{1-\frac{\alpha}{2}}^{[n-1]}\right) \qquad [10,21\text{-b}]$$

علماً أنَّ $t_k^{[n]}$ هي تلك النقطة على المحور الحقيقي التي يقع على يسارها وتحت منحنى الكثافة لتوزيع ستودنت (أو توزيع $t$) ذي الـ $n$ درجة حرية مساحة قدرها $k$، وأمّا $s$ فهي قيمة الإحصاءة $S$.

#### (١٠,٣,٢,٢) ملاحظات

١- في الحالة الخاصّة عندما يكون $\sigma$ مجهولاً وحجم العيِّنة $n$ كبيراً (أي إنَّ $32 \le n$) فإنَّ بعض المراجع تقوم باستخدام الفترة الآتية:

$$\left(\overline{x} - \frac{s}{\sqrt{n}}\,\mathfrak{z}_{1-\frac{\alpha}{2}}\ ,\ \overline{x} + \frac{s}{\sqrt{n}}\,\mathfrak{z}_{1-\frac{\alpha}{2}}\right) \qquad [10,21\text{-c}]$$

على أنّها $100\,(1-\alpha)\,\%$ فترة ثقة لـ $\mu$ بدلاً من الفترة [10-21-b]، ويبرِّرون ذلك من خلال كون الفوارق الناتجة عن استخدام توزيع ستودنت والتوزيع الطبيعي المعّياري من أجل قيم $32 \le n$ صغيرة، وتزداد صغراً كلّما كَبر حجم العيِّنة لدرجة أنّه يمكن معها النظر إلى هذه الفوارق كمقادير مهملة (أمعن النظر في الجدولين الخاصين بالتوزيعين المذكورين في نهاية الكتاب)، ويُضاف إلى ما سبق أنَّ الكثير من جداول توزيع ستودنت تقدِّم قيم $t_k^{[n]}$ من أجل بعض القيم لـ $n$ عندما تصبح $30 < n$ حيث تقدِّم تلك القيم على قفزات بتزايدات مختلفة 5 أو 10 أو...، ولكن في كتابنا هذا لن نعمل بهذه الملاحظة ونستخدم كلاً حسب توزيعه الأصل، والجدول الذي سيقدَّم يتضمن قيم $n$ بدءًا من 1 وحتى 100 بشكل متتال.

٢- إنَّ الفترة المُعطاة من خلال [10,21-a] صحيحة ما دام المجتمع الذي قيد الدراسة غير محدود، أو أنَّ المعاينة تتمُّ مع الإعادة من أجل المجتمعات المحدودة، وأمّا إذا كان هذا المجتمع محدوداً وحجمه $N$، وكانت المعاينة تتم دون إعادة، فعندئذ يُفضل استخدام الفترة الآتية عوضاً عن الفترة المقدَّمة في [10,21-a] :

$$\left(\overline{x}-\frac{\sigma_o}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\cdot z_{1-\frac{\alpha}{2}}\ ,\ \overline{x}+\frac{\sigma_o}{\sqrt{n}}\sqrt{\frac{N-n}{N-1}}\cdot z_{1-\frac{\alpha}{2}}\right) \qquad [10,21\text{-d}]$$

وتكون هذه الفترة أكثر دقَّةً عندما يكون حجم العيّنة $0.05\ N < n$ .

٣- فيما بعد وعندما نذكر عبارة **"الفترة [ ? , ? ]"** إنَّما نقصد بذلك الفترة ذات الرقم [ ? , ? ] . كذلك وعلى سبيل التذكير، عندما لا يُذكر شيء حول حجم المجتمع المدروس، فإنَّما نعني بذلك أنَّ حجم المجتمع غير محدود، وكذلك عندما لا يُذكر شيء حول نوع عمليّة المعاينة، فإنَّما يُقصد بذلك أنَّ المعاينة تتم مع الإعادة ما دام المجتمع محدوداً (**ذلك أنَّه في حال المجتمعات غير المحدودة، لا أهمية تذكر لنوع عملية المعاينة، سواء أكانت هذه العمليَّة تتم مع الإعادة أم بدون إعادة**).

## (١٠,٣,٢,٣) نتائج

من المبرهنة (١٠,٣,٢,١) والملاحظات (١٠,٣,٢,٢) السّابقتين نحصل على النتائج الآتية:

١- إذا كان الانحراف المعياري للمجتمع معلوماً $\sigma = \sigma_o$، فعندئذ:

أ- إذا كان حجم المجتمع **غير محدود** فإنَّ تقديرنا لمتوسط المجتمع $\mu$ بالقيمة $\overline{x}$ (قيمة $\overline{X}$) وعند مستوى من الثّقة قدره $1-\alpha$ سيترتب عليه خطأ أعظمي (سنرمز له بـ $er$) يساوي:

$$er = \pm\frac{\sigma_o}{\sqrt{n}}\cdot z_{1-\frac{\alpha}{2}} \qquad [10,22\text{-a}]$$

ومن أجل قيمة مُعطاة للخطأ $er = er_o$ يُقدَّر حجم العيِّنة التي يجب سحبها بحيث لا يتجاوز الخطأ المرتكب في تقديرنا لـ $\mu$ بالقيمة $\overline{x}$ أكثر من $er_o$ بوساطة المتباينة الآتية:

$$n \ \geq\ \left(\frac{\sigma_o}{er_o}\cdot z_{1-\frac{\alpha}{2}}\right)^2 \qquad [10,22\text{-b}]$$

ب- إذا كان حجم المجتمع **محدوداً** ويساوي $N$ ، فإنَّ تقديرنا لمتوسط المجتمع $\mu$ بالقيمة $\overline{x}$ وعند مستوى من الثّقة قدره $1-\alpha$ سيترتب عليه خطأ أعظمي (سنرمز له بـ $er$ أيضاً) يساوي:

$$er = \pm\frac{\sigma_o}{\sqrt{n}}\cdot\sqrt{\frac{N-n}{N-1}}\cdot z_{1-\frac{\alpha}{2}} \qquad [10,23\text{-a}]$$

ومن أجل قيمة مُعطاة للخطأ $er$ يُقدَّر حجم العيِّنة التي يجب سحبها بحيث لا يتجاوز الخطأ المرتكب في تقديرنا لـ $\mu$ بالقيمة $\overline{x}$ أكثر من $er_o$ بوساطة المتباينة الآتية:

$$n \ \geq\ N\cdot\sigma_o^2\cdot z^2_{1-\frac{\alpha}{2}}\cdot\left[(N-1)\,er_o^2+\sigma_o^2\cdot z^2_{1-\frac{\alpha}{2}}\right]^{-1} \qquad [10,23\text{-b}]$$

إنَّ الخطأ $er$ (وفي كلا الحالتين) يُدعى **الخطأ المحتمل** ، وذلك لأنَّه يكون مُحقَّقاً باحتمال معلوم يساوي $1-\alpha$ .

٢ - إذا كان الانحراف المعياري للمجتمع مجهولاً $\sigma = ?$، فعندئذ تقديرنا لمتوسط المجتمع $\mu$ بالقيمة $\overline{x}$ وعند مستوى من الثّقة قدره $1-\alpha$ سيترتب عليه خطأ أعظمي $er$ مقداره يساوي:

$$er = \pm \frac{s}{\sqrt{n}} \cdot t_{1-\frac{\alpha}{2}}^{[n-1]} \qquad [10,24]$$

ومن أجل قيمة $er_o$ مُعطاة للخطأ، فإنَّه يُقدَّر حجم العيِّنة التي يجب سحبها بحيث لا يتجاوز الخطأ المرتكب في تقديرنا ل $\mu$ بالقيمة $\overline{x}$ أكثر من $er_o$ بوساطة المتباينة الآتية:

$$n \;\geq\; \left( \frac{s}{er_o} \cdot t_{1-\frac{\alpha}{2}}^{[n-1]} \right)^2 \qquad [10,25]$$

تقدِّم لنا المبرهنة الآتية (التي سنقبلها دون برهان) فترة الثّقة ل $\sigma^2$ (تباين مجتمع طبيعي).

### (١٠,٣,٢,٤) مبرهنة

تحت الفرضيات المذكورة سابقاً في أولاً من (١٠,٣,٢) تُعطى $100(1-\alpha)\%$ فترة ثقة لـ $\sigma^2$ من خلال الفترة الآتية:

$$\left( (n-1)\, s^2 \Big/ \chi_{1-\frac{\alpha}{2}}^{[n-1]} \;,\; (n-1)\, s^2 \Big/ \chi_{\frac{\alpha}{2}}^{[n-1]} \right) \qquad [10,26]$$

علماً أنَّ $\chi_k^{[m]}$ هي تلك النقطة على المحور الحقيقي التي يقع على يسارها وتحت منحنى دالَّة الكثافة الاحتمالية لتوزيع كاي مربع ذي الـ $m$ درجة حرية مساحة قدرها $k$.

### (١٠,٣,٢,٥) ملاحظة

إنَّ معظم جداول توزيع كاي مربع تُعطى من أجل بعض القيم لـ $n$، وعندما تصبح قيمة $n$ أكبر 25 أو30 فإنَّها تقدَّم بتزايدات 5 أو 10 أو أكثر من ذلك، ولذلك فإنَّه من أجل قيم $n$ غير الموجودة في هذه الجداول يمكن استخدام جدول التوزيع الطبيعي عوضاً عن جدول توزيع كاي مربع من خلال حساب $\chi_\alpha^{[n]}$ وفقاً للعلاقة الآتية:

$$\chi_\alpha^{[n]} = \frac{1}{2}\left( z_\alpha + \sqrt{2n-1} \right)^2 \qquad [10,27]$$

حيث نحصل على قيمة تقريبية مقبولة في مثل هذه الحالات.

### (١٠,٣,٢,٦) أمثلة

١ - زعم أحد المتاجر أنَّ متوسط أرباحه اليومية على مرور أيام السنة يساوي 1000 وحدة نقدية بانحراف معياري قدره 100 وحدة نقدية وذلك عند مستوى من الثّقة قدره 95%. للتأكد من صحة زعم صاحب هذا المتجر قمنا بحساب الأرباح اليومية ل 25 يوماً سُحِبت عشوائياً من سجل مبيعات المتجر لسنة كاملة، فوجدنا المُعطيات الآتية:

الجدول (١٠,١)

| اليوم | المبلغ | اليوم | المبلغ | اليوم | المبلغ | اليوم | المبلغ |
|---|---|---|---|---|---|---|---|
| 1 | 900 | 8 | 1000 | 15 | 950 | 22 | 950 |
| 2 | 1000 | 9 | 850 | 16 | 1200 | 23 | 1010 |
| 3 | 950 | 10 | 800 | 17 | 1000 | 24 | 1000 |
| 4 | 1300 | 11 | 1000 | 18 | 1100 | 25 | 850 |
| 5 | 1000 | 12 | 950 | 19 | 1000 | | |
| 6 | 1010 | 13 | 1000 | 20 | 1010 | | |
| 7 | 1010 | 14 | 1100 | 21 | 800 | | |

ومن هذه البيانات نجد أنَّ $\overline{x} = 989.6$ و $s = 110.434$ ، ومن ثمَّ بفرض أنَّ الأرباح اليومية للمتجر تخضع للتوزيع الطبيعي على وجه التقريب بمتوسط $\mu$ وانحراف معياري $\sigma$ ، فعندئذ:

أ- إذا افترضنا أنَّ قيمة الانحراف المعياري التي قدَّمها المتجر موثوق بها (أي إنَّ $\sigma_o = 100 = \sigma$ معلومة) فإنَّ $100(1-\alpha)\%$ فترة ثقة لمتوسط أرباح المتجر تُعطى من خلال الفترة [10,21-d] أن حجم المجتمع الذي سُحبت منه محدوداً ويساوي $N = 365$ يوماً، حيث لدينا $1-\alpha = 0.95$ و $n = 25$ وكذلك باستخدام جدول التوزيع الطبيعي نجد:

$$z_{1-\frac{\alpha}{2}} = z_{1-\frac{0.05}{2}} = z_{0.975} = 1.96$$

وبعد التعويض في الفترة [10–21-d] كلاً بما يساويه نجد أنَّ 95% فترة ثقة لـ $\mu$ متوسط أرباح المتجر هي $(952.99 \,,\, 1026.22)$، وبما أنَّ القيمة (**1000 وحدة نقدية**) التي قدَّمها صاحب المتجر تقع في فترة الثّقة الموافقة لهذه الحالة، فإنَّنا سنقبل بادّعاء صاحب المتجر حول متوسط أرباحه اليومية عند مستوى الثّقة 95%، وفي هذه الحالة تكون قيمة الخطأ المرتكب في تقديرنا لمتوسط أرباح المتجر بقيمة $\overline{x} = 989.6$ في حال كانت قيمة الانحراف المعياري المقدَّمة معلومة تساوي:

$$er = \pm \frac{\sigma_o}{\sqrt{n}} \sqrt{\frac{N-n}{N-1}} \cdot z_{1-\frac{\alpha}{2}} = \pm \frac{100}{\sqrt{25}} \sqrt{\frac{365-25}{365-1}} \cdot 1.96 = \pm 36.62$$

ومن أجل حساب حجم العيِّنة التي يجب سحبها (عدد الأيام التي يجب حساب متوسط أرباحها) بحيث إنَّ الخطأ المرتكب في تقديرنا لـ $\mu$ بالقيمة $\overline{x} = 989.6$ لا يتجاوز $er_o = 35$ وحدة نقدية لدينا:

$$n \geq \frac{N \cdot \sigma_o^2 \cdot z_{1-\frac{\alpha}{2}}^2}{(N-1)\, er_o^2 + \sigma_o^2 \cdot z_{1-\frac{\alpha}{2}}^2} = \frac{365\,(100)(1.96)^2}{(364)(35)^2 + (100)^2\,(1.96)^2} = 28.95$$

وبما أنَّ حجم العيِّنة $n$ يجب أن يكون عدداً صحيحاً، فإنَّه علينا أخذ $n = 29$ على الأقل.

أمّا إذا افترضنا أنَّ حجم المجتمع غير محدود (بمعنى أن المتجر يعمل منذ زمن بعيد جداً وأنَّ العيّنة التي سُحبت كانت من دفاتر سجلاته السابقة منذ نشأة المتجر)، فعندئذ نعيِّن 95% فترة ثقة لـ $\mu$ متوسط أرباح من خلال الفترة [10–21-a]، وبعد التعويض وتنفيذ الحسابات المطلوبة نجدها تساوي:

$$\left(\overline{x} - \frac{\sigma_o}{\sqrt{n}} z_{1-\frac{\alpha}{2}} \,,\, \overline{x} + \frac{\sigma_o}{\sqrt{n}} z_{1-\frac{\alpha}{2}}\right) = \left(989.6 - \frac{100}{\sqrt{25}} 1.96 \,,\, 989.6 + \frac{100}{\sqrt{25}} 1.96\right) = (950.40 \,,\, 1028.80)$$

ومن ثمَّ لا يطرأ أي تغيير على القرار الذي اتخذناه سابقاً عند المستوى الثّقة 95%، وفي هذه الحالة تكون قيمة الخطأ المرتكب في تقديرنا لمتوسط أرباح المتجر بقيمة $\overline{x} = 989.6$ في حال كانت قيمة الانحراف المعياري المقدَّمة معلومة تساوي:

$$er = \pm \frac{\sigma_o}{\sqrt{n}} \cdot z_{1-\frac{\alpha}{2}} = \pm \frac{100}{\sqrt{25}} \cdot 1.96 = \pm 39.20$$

وأمّا إذا أردنا حساب حجم العيِّنة التي يجب سحبها بحيث لا يتجاوز الخطأ المرتكب في تقديرنا لـ $\mu$ بالقيمة $\overline{x} = 989.6$ بأكثر من 35 وحدة نقدية، فإنَّنا نجده يساوي:

$$n \geq \left(\frac{\sigma_o}{er_o} \cdot z_{1-\frac{\alpha}{2}}\right)^2 = \left(\frac{100}{35} \cdot 1.96\right)^2 = 31.36$$

وهذا يعني أنَّه في هذه الحالة يجب أخذ $n = 32$ على الأقل.

ب- إذا افترضنا أنَّ قيمة الانحراف المعياري التي قدَّمها المتجر غير موثوق بها (أي إنَّ $\sigma = ?$ مجهولة) فإنَّ $100(1-\alpha)\%$ فترة ثقة لمتوسط أرباح المتجر تُعطى من خلال الفترة [10,21-b]، وبما أنَّه لدينا $n = 25$ و $1-\alpha = 0.95$ و $t_{1-\frac{\alpha}{2}}^{[n-1]} = t_{0.975}^{[24]} = 2.064$ (من جدول توزيع ستودنت)، فإنَّه سيكون لدينا 95% فترة ثقة لـ $\mu$ متوسط أرباح المتجر لها العرض الآتي:

$$\left(\bar{x} - \frac{s}{\sqrt{n}}\, t_{1-\frac{\alpha}{2}}^{[n-1]} \,,\, \bar{x} + \frac{s}{\sqrt{n}}\, t_{1-\frac{\alpha}{2}}^{[n-1]}\right) = \left(989.6 - \frac{110.434}{\sqrt{25}}\ 2.064 \,,\, 989.6 + \frac{110.434}{\sqrt{25}}\ 2.064\right)$$
$$= (944.01 \,,\, 1035.19)$$

وفي هذه الحالة نلاحظ أنَّ القرار الذي اتخذناه سابقاً يبقى صحيحاً عند المستوى الثّقة 95% أيضاً، وتكون قيمة الخطأ المرتكب في تقديرنا لمتوسط أرباح المتجر بقيمة $\bar{x} = 989.6$ في حال كانت قيمة الانحراف المعياري المقدَّمة مجهولة تساوي:

$$er = \pm \frac{s}{\sqrt{n}}\, t_{1-\frac{\alpha}{2}}^{[n-1]} = \pm \frac{110.434}{\sqrt{25}}\ 2.064 = \pm 45.59$$

وأمَّا إذا أردنا حساب حجم العيِّنة التي يجب سحبها (عدد الأيام التي يجب حسابها متوسط أرباحها) بحيث لا يتجاوز الخطأ المرتكب في تقديرنا لـ $\mu$ بـ $\bar{x} = 989.6$ قيمة معلومة تساوي $er_o = 35$ وحدة نقدية فإنَّنا نجده يساوي:

$$n \ \geq \ \left(\frac{s}{er_o} \cdot t_{1-\frac{\alpha}{2}}^{[n-1]}\right)^2 = \left(\frac{110.43}{35} \cdot 2.064\right)^2 = 42.41$$

وهذا يعني أنَّه يجب أخذ $n = 43$ على الأقل في هذه الحالة.

هـ- أخيراً لنقم بتعيين 95% فترة ثقة ل $\sigma^2$ تباين مجتمع الأرباح اليومية لمتجر، حيث نعلم أنَّ $100(1-\alpha)\%$ فترة ثقة لـ $\sigma^2$ تُعطى من خلال الفترة [10–26]، حيث لدينا:

$$\chi_{1-\frac{\alpha}{2}}^{[m]} = \chi_{0.975}^{[24]} = 39.364 \qquad \& \qquad \chi_{\frac{\alpha}{2}}^{[m]} = \chi_{0.025}^{[24]} = 12.401$$

ومن ثمّ ينتج لدينا أنَّ $100(1-\alpha)\%$ فترة ثقة لـ $\sigma^2$ هي:

$$\left(\frac{(n-1)\,s^2}{\chi_{1-\frac{\alpha}{2}}^{[n-1]}} \,,\, \frac{(n-1)\,s^2}{\chi_{\frac{\alpha}{2}}^{[n-1]}}\right) = \left(\frac{24(110.43)^2}{39.364} \,,\, \frac{16(110.43)^2}{12.401}\right) = (7435.09 \,,\, 23600.91)$$

### (١٠,٣,٣) فترة الثّقة لمَعْلمة مجتمع برنولي

سنفترض على مدى هذه الفقرة أنَّه لدينا $[\Omega \,,\, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع لتوزيع برنولي بمَعْلمة $0 < p < 1$، وأنَّ $\mathcal{X}_n = (X_1, X_2, ..., X_n)$ عيِّنة من هذا المجتمع. عندئذ تفيدنا مبرهنة النهاية المركزية لموافير-لابلاس أنَّه سيكون للمتغيِّر العشوائي:

$$B := \frac{(\overline{\pi} - p)}{\sqrt{\frac{p\,(1-p)}{n}}} = \sqrt{\frac{n}{p\,(1-p)}}\ \ (\overline{\pi} - p) \qquad [10,27]$$

توزيع طبيعي معياري $\mathbf{N}(0\,,1)$ ، علماً أنَّ $\overline{\pi}=\frac{1}{n}\sum_{i=1}^{n}X_i$ يمثّل التكرار النسبي للنجاحات في العيِّنة وهو مقدِّر غير منحاز لمَعْلمة المجتمع $p$ ، لأنَّ $\mathbf{E}\,\overline{\pi}=p$ ، وكذلك لدينا $\mathbf{var}\,\overline{\pi}=\frac{p\,(1-p)}{n}$ .

إنَّ المبرهنة الآتية (سنقبلها دون برهان) تُعطينا $100\,(1-\alpha)\,\%$ فترة ثقة لمَعْلمة المجتمع البرنولي.

**(١٠,٣,٣,١) مبرهنة**

تحت الفرضيات المذكورة آنفًا، تُعطى $100\,(1-\alpha)\,\%$ فترة ثقة للمَعْلمة $p$ من خلال الفترة الآتية:

$$\left(\frac{\bar{p}+\frac{z^2_{1-\frac{\alpha}{2}}}{2n}-\frac{z_{1-\frac{\alpha}{2}}}{\sqrt{n}}\sqrt{\bar{p}\,(1-\bar{p})+\frac{z^2_{1-\frac{\alpha}{2}}}{4n}}}{1+\frac{z^2_{1-\frac{\alpha}{2}}}{n}}\;,\;\frac{\bar{p}+\frac{z^2_{1-\frac{\alpha}{2}}}{2n}+\frac{z_{1-\frac{\alpha}{2}}}{\sqrt{n}}\sqrt{\bar{p}\,(1-\bar{p})+\frac{z^2_{1-\frac{\alpha}{2}}}{4n}}}{1+\frac{z^2_{1-\frac{\alpha}{2}}}{n}}\right)$$

علماً أنَّ $\bar{p}$ هي قيمة الإحصاءة $\overline{\pi}$ .

**(١٠,٣,٣,٢) ملاحظات**

١- من الناحية التطبيقية يُمكننا إهمال قيم الحدود $\frac{z^2_{1-\frac{\alpha}{2}}}{n}$ و $\frac{z^2_{1-\frac{\alpha}{2}}}{2n}$ و $\frac{z^2_{1-\frac{\alpha}{2}}}{4n}$ في الفترة السابقة وذلك بسبب صغَر المقدار $z_{1-\frac{\alpha}{2}}$ بالمقارنة مع قيم $n$ التي تُؤخَذ كبيرة بقدر كاف في معظم الدراسات الإحصائية، وذلك لأنَّه في حالة كانت قيمة $z_{1-\frac{\alpha}{2}}$ ليست صغيرة بما فيه الكفاية بالنسبة إلى قيمة $n$ فإنَّ تقريب موافير-لابلاس يصبح غير ممكن في تعيين فترة الثّقة الأخيرة، ولذلك سنستخدم الفترة الآتية عوضاً عن الفترة السابقة على أنَّها $100\,(1-\alpha)\,\%$ فترة ثقة للمَعْلمة $p$ :

$$\left(\bar{p}-z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{p}\,(1-\bar{p})}{n}}\;,\;\bar{p}+z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{p}\,(1-\bar{p})}{n}}\right) \qquad [10,28]$$

وتزداد دقّة استخدام هذه الفترة كلّما كَبُرَت قيمة $n$ ، وعلى العموم يمكننا قبول هذه الفترة عندما يكون $n\geq 32$ مع تحقُّق المتباينتين $5<n\,p$ و $5<n\,(1-p)$ أيضاً.

٢- نلاحظ هنا أنَّ تقديرنا لمَعْلمة المجتمع $p$ بالقيمة $\bar{p}$ سيرافقه خطأ أعظمي $er$ مقداره يساوي:

$$er=\pm\, z_{1-\frac{\alpha}{2}}\sqrt{\frac{\bar{p}\,(1-\bar{p})}{n}} \qquad [10,29]$$

٣- يمكننا تقدير حجم العيِّنة $n$ التي يجب سحبها من المجتمع بحيث لا يتجاوز الخطأ المرتكب في تقديرنا للمَعْلمة $p$ بالقيمة $\bar{p}$ بأكثر من $er$، وذلك باستخدام المتباينة الآتية:

$$n\geq\left(\frac{z_{1-\frac{\alpha}{2}}}{er}\right)^2\bar{p}\,(1-\bar{p}) \qquad [10,30]$$

**(٣,٣,٣,١٠) أمثلة**

١ – تَدَّعي مؤسسة تعليمية أنَّها على ثقة قدرها % 98 بأنَّ نسبة الطلبة المتعثرين لديها لا تتجاوز 0.05، فأرادت دائرة الجودة (التي تتبعها هذه المؤسسة) التّحقّق من صحة ادعاء هذه المؤسسة التعليمية، فقامت بسحب عيِّنة من طلاب هذه المؤسسة بحجم 256 طالباً، وبحثت في سجلّ كل طالب من هذه العيّنة لدى المؤسسة فوجدت 16 طالباً متعثراً، فهل تشير هذه النتيجة إلى صحة ادعاء المؤسسة التعليمية عند مستوى الثّقة الذي قدَّمته هذه المؤسسة؟

الحـل: بملاحظة أنَّ تفحص سجل كل طالب هي تجربة برنولية، فإنَّ ذلك يعني أنّنا أمام مجتمع برنولي (ذلك أنَّ أيّ طالب إمّا أن يكون متعثّراً أو غير متعثّر) بمَعْلمة $p = 0.05$ (هنا الحصول على طالب متعثّر يُعدّ نجاحاً)، ولكن بسبب الشّك في قيمة هذه المَعْلمة فإنّنا سنتعامل معها على أنّها مجهولة، ولذلك فمن أجل الإجابة على السؤال الطروح علينا تعيين %98 فترة ثقة للمَعْلمة $p$، وبما أنَّ حجم العيّنة كبير وتفحص سجلات الطلاب هو تكرارات مستقلّة عن بعضها، فعندئذ تُعطى فترة الثقة المطلوب تعيينها بوساطة [28–10]، حيث لدينا من معطيات المسألة ما يلي:

$$\bar{p} = \frac{16}{256} = 0.0625 \quad \& \quad z_{1-\frac{\alpha}{2}} = z_{0.99} = 2.33$$

ومن ثّم تكون لدينا %98 فترة ثقة للمَعْلمة $p$ هي:

$$\left( 0.0625 - 2.33\sqrt{\frac{0.0625\ (1-0.0625)}{16}} \ , \ 0.0625 + 2.33\sqrt{\frac{0.0625\ (1-0.0625)}{16}} \right)$$

وبعد تنفيذ الحسابات نحصل على الفترة ( 0.0785 – , 0.2035 )، ومن ثّم نقبل بادعاء هذه المؤسسة التعليمية عند مستوى الثّقة 0.98 لأنَّ نسبة الطلبة المتعثرين التي قدَّمتها 0.05 تنتمي إلى هذه الفترة.

٢ – نقوم بدراسة ميدانية حول تأثير نوع من الأسمدة على إنتاج محصول زراعي، وأردنا التّحقُّق من نتيجة هذه التجربة بحيث إنه عند مستوى من الثّقة قدره %95 يجب أن تكون الفوارق التي سنحصل عليها في حدود 0.05± من النسبة الحقيقية لنجاح هذه التجربة، وهذا يعني أنَّ الخطأ المرتكب في تقدير $p$ بالتكرار النسبي للنجاحات في العيّنة $\bar{p}$ يجب أن يكون في حدود 0.05± . فإذا كانت التجارب الخاصّة بهذه النتيجة تُكرَّر بشكل مستقل بعضها عن بعض وتحت الشروط نفسها، فعندئذ كم مرَّة يجب أنَّ نُكرِّر التجربة عند مستوى ثقة قدره %95 لكي يكون الفارق في حدود 0.15± من النسبة الحقيقية لنجاح هذه التجربة، وذلك من أجل الحالتين الآتيتين:

أ – إذا علمنا من باحثين آخرين أنَّ نسبة النجاح لهذه التجربة هو 0.10 ؟

ب–إذا كنّا لا نعلم شيئاً عن نسبة النجاح لهذه التجربة؟

الحـل: للإجابة على هذه الأسئلة لدينا من أجل الطلب:

أ– نجد من معطيات المسألة أنَّ:

$$er = \pm 0.15 \quad \& \quad \bar{p} = 0.10 \quad \& \quad z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$$

ومن ثّم عدد التجارب التي يجب تكرارها تُقدَّر بالمتباينة [30–10]، وبالتعويض في تلك المتباينة نجد أنَّ:

$$n \geq \left(\frac{1.96}{0.15}\right)^2 (0.10)\,(0.90) = 15.3664$$

ومن ثّم يجب تكرار هذه التجربة 16 مرَّة على الأقل حتى يصبح الفارق في حدود 0.15± من النسبة الحقيقية لنجاح هذه التجربة.

ب- نجد من معطيات المسألة أنَّ:

$$er = \pm 0.15 \quad \& \quad \overline{p} = ? \quad \& \quad z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$$

ونلاحظ أنَّه لا بدَّ من تعيين قيمة $\overline{p}$ أولاً. إنَّ ذلك يتمّ من خلال حساب قيمة $\overline{p}$ التي يكون فيها للدالَّة $f(\overline{p}) = \overline{p}\,(1-\overline{p})$ نهاية عظمى، وفي هذا الصدد يمكن التّحقُّق بسهولة من أنَّ $f(\overline{p})$ يكون لها نهاية عظمى عندما يكون $\overline{p} = \frac{1}{2}$ وذلك من خلال وضع $f'(\overline{p}) = 0$، ومن ثمّ تكون أكبر قيمة لـ $n$ التي من أجلها نحافظ على الدقّة المطلوبة (المتمثّلة بالخطأ $er$) آنفًا وبغض النظر عن القيمة الفعليَّة لنسبة نجاح التجربة $\overline{p}$ تعطى وفقاً للمتباينة [30–10] بالقيمة $n \leq 42.684$، وهذا يعني أنَّه في هذه الحالة يجب تكرار هذه التجربة 43 مرَّة على الأقل.

### (١٠,٣,٤) فترات الثّقة للفرق بين متوسطي مجتمعين طبيعيين مستقلَّين ولنسبة تشتتهما

سنفرض على مدى هذه الفقرة أنَّه لدينا مجتمعين طبيعيين مستقلَّين. في مثل هذه الحالات يُناقش تعيين فترات الثّقة الخاصة بالفرق بين متوسطي هذين المجتمعين، وكذلك تعيين فترات الثّقة الخاصة بنسبة التباين لهما.

ليكن $[\Omega_1, \mathbb{S}_1, \mathcal{P}_1]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع للتوزيع الطبيعي $\mathbf{N}(\mu_1, \sigma_1)$، و $\mathcal{X}_n = (X_1, X_2, ..., X_n)$ عيِّنة مسحوبة من هذا المجتمع بمتوسط $\overline{x}$ وانحراف معياري $s_X$. كذلك ليكن $[\Omega_2, \mathbb{S}_2, \mathcal{P}_2]$ مجتمع آخر مستقل عن المجتمع السّابق، وموصوف من خلال متغيِّر عشوائي $Y$ خاضع للتوزيع الطبيعي $\mathbf{N}(\mu_2, \sigma_2)$، ولنأخذ $\mathcal{Y}_m = (Y_1, Y_2, ..., Y_m)$ عيِّنة مسحوبة من هذا المجتمع الأخير بمتوسط $\overline{y}$ وانحراف معياري $s_Y$.

تقدِّم لنا المبرهنة الآتية (التي سنقبلها دون برهان) فترات الثّقة للفرق بين متوسطي مجتمعين طبيعيين.

#### (١٠,٣,٤,١) مبرهنة

تحت الفرضيات المذكورة آنفًا، فإنَّه لإعطاء $100\,(1-\alpha)\,\%$ فترة ثقة للفرق $\mu_1 - \mu_2$ سنناقش الحالات الآتية:

١- إذا كان الانحراف المعياري لكلٍّ من المجتمعين معلوماً بحيث $\sigma_1 = \sigma_{1,o}$ و $\sigma_2 = \sigma_{2,o}$، فإنَّ $100\,(1-\alpha)\,\%$ فترة ثقة للفرق $\mu_1 - \mu_2$ تُعطى من خلال الفترة الآتية:

$$\left( (\overline{x} - \overline{y}) - z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_{1,o}^2}{n} + \frac{\sigma_{2,o}^2}{m}} \;\; , \;\; (\overline{x} - \overline{y}) + z_{1-\frac{\alpha}{2}} \sqrt{\frac{\sigma_{1,o}^2}{n} + \frac{\sigma_{2,o}^2}{m}} \right) \qquad [10,31]$$

٢- إذا كان الانحراف المعياري لكلٍّ من المجتمعين مجهولاً $\sigma_1 = ?$ و $\sigma_2 = ?$ ولكن مع الافتراض بتساويهما (أي أنَّ $\sigma_1 = \sigma_2$)، فعندئذ تُعطى $100\,(1-\alpha)\,\%$ فترة ثقة للفرق $\mu_1 - \mu_2$ من خلال الفترة التي لحديها الأدنى (الموافق لإشارة -) والأعلى (الموافق لإشارة +) العرض الآتي على الترتيب:

$$(\overline{x} - \overline{y}) \mp t_{1-\frac{\alpha}{2}}^{[n+m-2]} \sqrt{\frac{(n+m)\left[(n-1)s_X^2 + (m-1)s_Y^2\right]}{n\,m\,(n+m-2)}} \qquad [10,32]$$

٣- إذا كان الانحراف المعياري لكلٍّ من المجتمعين مجهولاً $\sigma_1 = ?$ و $\sigma_2 = ?$ ولكن مع الافتراض بعدم تساويهما (أي إنَّ $\sigma_1 \neq \sigma_2$)، فإنَّ دراسة هذه الحالة ليست بسيطة كما ذكرنا ذلك في الفصل السابق، حيث سيكون للمتغيِّر العشوائي $T^*$ الآتي (قدِّم في الفصل السابق بالعلاقة رقم [9–28]):

$$T^* = \frac{(\overline{X} - \overline{Y}) - (\mu_1 - \mu_2)}{\sqrt{\left(S_X^2 / n\right) + \left(S_Y^2 / m\right)}}$$

توزيع ستودنت على وجه التقريب، وحينئذ تُعطى $\%(1-\alpha)\,100$ فترة ثقة تقريبية للفرق $\mu_1-\mu_2$ من خلال الفترة الآتية:

$$\left((\bar{x}-\bar{y})-t_{1-\frac{\alpha}{2}}^{[\zeta]}\sqrt{\frac{(n-1)s_X^2+(m-1)s_Y^2}{n\cdot m}}\ ,\ (\bar{x}-\bar{y})+t_{1-\frac{\alpha}{2}}^{[\zeta]}\sqrt{\frac{(n-1)s_X^2+(m-1)s_Y^2}{n\cdot m}}\right) \qquad [10,33]$$

علماً أنَّ عدد درجات الحرية $\nu$ الموافقة لهذه الحالة تحسب من العلاقة المقدّمة جانباً (تقريب ساتيرثويت)، وأنَّ $\lceil x\rceil$ هو أصغر عدد صحيح أكبر من $x$.

$$\nu=\left\lceil\frac{\left(\frac{s_X^2}{n}+\frac{s_Y^2}{m}\right)^2}{\sqrt{\left[\left(s_X^2/n\right)^2\Big/(n-1)\right]+\left[\left(s_Y^2/n\right)^2\Big/(m-1)\right]}}\right\rceil$$

**(٢, ٤, ٣, ١٠) ملاحظة**

من فترات الثّقة السابقة [31–10] و [32–10] و[33–10]

نلاحظ ما يلي:

١- إذا كان الطرف الأيمن **سالباً** لأيّة فترة من الفترات الثلاث السابقة، فعندئذ يكون لدينا $\mu_1<\mu_2$ باحتمال قدره $1-\alpha$.

٢- إذا كان الطرف الأيسر **موجباً** لأيّة فترة من الفترات الثلاث السابقة، فعندئذ يكون لدينا $\mu_1>\mu_2$ باحتمال قدره $1-\alpha$.

٣- إذا كان الطرف الأيمن **موجباً** والطرف الأيسر **سالباً** لأيّة فترة من الفترات الثلاث السابقة، فعندئذ سيكون لدينا $\mu_1=\mu_2$ باحتمال قدره $1-\alpha$، وما يظهر من فروق في قيم $\bar{X}$ و $\bar{Y}$ إنّما هي فروق ظاهرية لا أهمّية لها ما دامت قيمة $1-\alpha$ قريبة من الواحد.

المبرهنة الآتية تعطينا $\%(1-\alpha)\,100$ فترة الثّقة لنسبة تبايني مجتمعين طبيعيين.

**(٣, ٤, ٣, ١٠) مبرهنة**

تحت الفرضيات المقدَّمة سابقاً في هذه الفقرة، فإنَّ $\%(1-\alpha)\,100$ فترة ثقة للنسبة $\frac{\sigma_1^2}{\sigma_2^2}$ تُعطى من خلال الفترة الآتية:

$$\left(\frac{S_X^2}{S_Y^2}.f_{\frac{\alpha}{2}}^{[m-1\,;\,n-1]}\ ,\ \frac{S_X^2}{S_Y^2}.f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]}\right) \qquad [10,34\text{-a}]$$

ولكن بسبب أنَّ الجدول المقدَّم في ملحق الجداول الإحصائية في نهاية هذا الكتاب لا يعطينا قيماً للحد $f_{\frac{\alpha}{2}}^{[m-1\,;\,n-1]}$ من أجل القيم الصغيرة لـ $\alpha$، ولذلك يُستعاض عن ذلك الحد بالمقدار $\left(f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]}\right)^{-1}$، حيث يُبرهن على أنَّ:

$$f_k^{[u\,;\,v]}=\left(f_{1-k}^{[u\,;\,v]}\right)^{-1} \qquad for\ \ u,v\in\mathbb{N}\ \ \&\ \ 0<k<1$$

ومن ثمَّ يصبح للفترة [34-a–10] العرض الآتي:

$$\left(\frac{S_X^2}{S_Y^2}\cdot\left(f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]}\right)^{-1}\ ,\ \frac{S_X^2}{S_Y^2}.f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]}\right) \qquad [10,34\text{-b}]$$

علماً أنَّ $f_k^{[u;v]}$ هي تلك النقطة على المحور الحقيقي التي يقع على يسارها وتحت منحنى الكثافة الاحتمالية لتوزيع **فيشر** ذي ال $u$ (**درجة حرية بسط**) و $v$ (**درجة حرية مقام**) مساحة قدرها $k$.

**(١٠,٣,٤,٤) نتائج**

من فترة الثّقة [10,34-a] أو [10,34-b] نستنتج ما يلي:

١- إذا كان الطرف الأيسر للفترة الأخيرة أكبر من الواحد، فعندئذ سيكون لدينا $\sigma_2^2 < \sigma_1^2$ باحتمال قدره $1-\alpha$ .

٢- إذا كان لدينا الطرف الأيمن للفترة الأخيرة أصغر من الواحد، فعندئذ سيكون $\sigma_2^2 > \sigma_1^2$ باحتمال قدره $1-\alpha$ .

٣- إذا كان لدينا الطرف الأيمن للفترة الأخيرة أكبر من الواحد وطرفها الأيسر أصغر من الواحد، فعندئذ سيكون لدينا $\sigma_2^2 = \sigma_1^2$ باحتمال قدره $1-\alpha$، ومن ثمّ باحتمال قدره $1-\alpha$ سيكون التباين للمجتمعين متساويين، وما يظهر من فوارق في قيم $S_X^2$ و $S_Y^2$ إنّما هي فوارق ظاهريّة لا أهميّة لها مادام مستوى الثّقة $1-\alpha$ قريب من الواحد.

**(١٠,٣,٤,٥) مثال**

تقدَّمت شركتان مستقلّتان بعضهما عن بعض لإنتاج الأطعمة المعلّبة (الكنسروة) بعروض بيع لمنتجها لإحدى المؤسسات المهتمّة بهذا النوع من التجارة. حيث ادّعت الشركة الأولى أنّها على ثقة قدرها 0.98 بأنّ متوسط وزن العلبة لديها 1000 غرام بانحراف معياري قدره 50 غرامًا، في حين ادّعت الشركة الثانية أنّها على ثقة قدرها 0.98 بأنّ متوسط وزن العلبة لديها 1050 غرامًا بانحراف معياري قدره 25 غرامًا. لكي تقرِّر المؤسسة أي الشركتين ستختار قامت بأخذ عيِّنة $\mathcal{X}_{81}$ من الشركة الأولى فوجدت أنَّ متوسط وزن العلبة في هذه العيّنة $\overline{x} = 1050$ غرامًا بانحراف معياري $s_X = 45$ غرامًا، وكذلك أخذت عيِّنة $\mathcal{Y}_{64}$ من الشركة الثانية فوجدت أنَّ متوسط وزن العلبة في هذه العيّنة $\overline{y} = 1025$ غرامًا بانحراف معياري $s_Y = 50$ غرامًا. الآن، لو افترضنا أنَّ أوزان العلب تخضع للتوزيع الطبيعي، فأي الشركتين يُفضَّل اختيارها من قبل المؤسسة عند مستوى من الثّقة 0.98؟

الحـل: للإجابة على هذا السؤال علينا القيام بتعيين 98% فترة ثقة للفرق $\mu_1-\mu_2$، وهنا علينا أن نناقش الحالات الآتية:

أ- إذا أخذنا بقيم الانحراف المعياري المقدَّمة من قبل الشركتين، وهذا يعني أنَّ التباين لكلٍّ من مجتمعي العلب معلوم حيث لدينا بخصوص العينة والمجتمع الأول $n = 81$، $\overline{x} = 1050$ و $\sigma_{1,o} = 50$، وأمّا بخصوص العينة والمجتمع الثاني فلدينا $m = 64$، $\overline{y} = 1025$ و $\sigma_{2,o} = 25$، وعلاوةً على ذلك لدينا $z_{1-\frac{\alpha}{2}} = z_{1-0.01} = z_{0.99} = 2.33$، ومن ثمَّ تُعطى 98 % فترة ثقة للفرق $\mu_1-\mu_2$ من خلال الفترة [31–10]، وبالتعويض عن كلّ رمز بما يساويه يكون لحدي هذه الفترة العرض الآتي:

$$\left(\overline{x}-\overline{y}\right) \mp z_{1-\frac{\alpha}{2}}\sqrt{\frac{\sigma_{1,o}^2}{n}+\frac{\sigma_{2,o}^2}{m}} = \left(1050-1025\right) \pm 2.33\sqrt{\frac{2500}{81}+\frac{625}{64}}$$

والتي تعطينا الفترة (10.15 , 39.85)، وبما أنَّ طرفي هذه الفترة موجب فإنَّ ذلك يعني أنَّ $\mu_1 > \mu_2$ باحتمال قدره 0.98، ومن ثمّ نفضِّل اختيار الشركة الأولى لأنَّ متوسط وزن العلب أكبر من مثيلاتها لدى الشركة الثانية.

ب- إذا تجاهلت المؤسسة قيم الانحراف المعياري المقدَّمة من قبل الشركتين، وهذا يعني أنَّ التباين لكلٍّ من مجتمعي العلب مجهول حيث لديها في هذه الحالة $\sigma_1 = ?$ و $\sigma_2 = ?$، وبما أنَّ التباين لكلٍّ من مجتمعي العلب مجهولاً فعلينا أولاً أن نبيِّن إن كان $\sigma_1^2 = \sigma_2^2$ أو لا.

من أجل ذلك علينا تعيين 98% فترة ثقة للنسبة $\dfrac{\sigma_1^2}{\sigma_2^2}$، علماً أنَّ 100 $(1-\alpha)$ % فترة ثقة لهذه النسبة تعطى من خلال الفترة [10–34-b]، حيث لدينا من جدول توزيع فيشر:

$$f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]} = f_{0.99}^{[63\,;\,80]} \approx f_{0.99}^{[60\,;\,60]} = 1.84$$

ومن ثمّ يكون للفترة المذكورة آنفًا العرض الآتي:

$$\left(\frac{s_X^2}{s_Y^2}\cdot\left(f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]}\right)^{-1}\ ,\ \frac{s_X^2}{s_Y^2}\cdot f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]}\right)=\left(\frac{2025}{2500}\cdot\frac{1}{1.84}\ ,\ \frac{2025}{2500}\cdot 1.84\right)=\left(0.440\ ,\ 1.490\right)$$

وبما أن طرفها الأيسر أصغر من الواحد، والأيمن أكبر من الواحد؛ فإنَّ ذلك يعني أنَّ $\sigma_1^2=\sigma_2^2$ عند مستوى الثّقة 98%، وبالتالي تُعطى 98 % فترة ثقة للفرق $\mu_1-\mu_2$ من خلال الفترة التي لحديها الأدنى والأعلى العرض [10–32] ، حيث لدينا من جدول توزيع ستودنت $t_{1-\frac{\alpha}{2}}^{[m+n-2]}=t_{0.99}^{[143]}=2.33$ ، وبالتعويض عن كلّ رمز بما يساويه في العرض [10–32] نجد لحدّي الفترة العرض الآتي:

$$\left(\overline{x}-\overline{y}\right)\mp t_{1-\frac{\alpha}{2}}^{[n+m-2]}\sqrt{\frac{(n+m)\left[(n-1)s_X^2+(m-1)s_Y^2\right]}{n\,m\,(n+m-2)}}=\left(1050-1025\right)\pm 2.33\sqrt{\frac{(81+64)\left[(80)\,2025+(63)\,2500\right]}{(81)(64)(143)}}$$

والتي تعطينا الفترة $(6.58\ ,\ 43.42)$ ، وفي هذه الحالة نجد أنَّ $\mu_1>\mu_2$ باحتمال قدره 0.98 أيضاً، ومن ثمّ نفضّل اختيار الشركة الأولى أيضاً.

### (١٠,٣,٥) فترة الثّقة للفرق بين مَعْلمتي مجتمعين برنوليين مستقلّين

ليكن $[\Omega_1\ ,\ \mathcal{S}_1\ ,\ \mathcal{P}_1]$ مجتمع برنولي بمَعْلمة $p_1$ ، ولنأخذ $\mathcal{X}_n=(X_1,X_2,\ldots,X_n)$ عيِّنة مسحوبة من هذا المجتمع. كذلك ليكن $[\Omega_2\ ,\ \mathcal{S}_2\ ,\ \mathcal{P}_2]$ مجتمعًا برنوليًا آخر مستقلّ عن المجتمع السابق بمَعْلمة $p_2$ ، ولنأخذ $\mathcal{Y}_m=(Y_1,Y_2,\ldots,Y_m)$ عيِّنة مسحوبة من هذا المجتمع الثاني. علماً أنَّ $n$ و $m$ قيم كبيرة بقدر كاف.

تُعطينا المبرهنة الآتية (التي سنقبلها دون برهان) $100\,(1-\alpha)\,\%$ فترة ثقة للفرق $p_1-p_2$ .

#### (١٠,٣,٥,١) مبرهنة

تحت الفرضيات المقدَّمة آنفًا تُعطى $100\,(1-\alpha)\,\%$ فترة ثقة (تقريبية) للفرق $p_1-p_2$ من خلال الفترة الآتية:

$$\left(\left(\overline{p}_1-\overline{p}_2\right)-z_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{p}_1(1-\overline{p}_1)}{n}+\frac{\overline{p}_2(1-\overline{p}_2)}{m}}\ \ ,\ \left(\overline{p}_1-\overline{p}_2\right)+z_{1-\frac{\alpha}{2}}\sqrt{\frac{\overline{p}_1(1-\overline{p}_1)}{n}+\frac{\overline{p}_2(1-\overline{p}_2)}{m}}\right)\qquad [10,34]$$

علماً أنَّ $p_1$ و $p_2$ هي قيمة الإحصاء $\overline{\pi}_1=\frac{1}{n}\sum_{k=1}^{n}X_k$ و $\overline{\pi}_2=\frac{1}{m}\sum_{k=1}^{m}Y_k$ الذي يمثّل التكرار النسبي للنجاحات في العيّنة $\mathcal{X}_n$ و $\mathcal{Y}_m$ على الترتيب.

#### (١٠,٣,٥,٢) ملاحظات

١ – إنَّ الفترة الأخيرة هي فترة تقريبية لـ $100\,(1-\alpha)\,\%$ فترة ثقة لـ $p_1-p_2$ ، وتزداد دقَّة هذه الفترة كلّما كَبُرَ حجم العيِّنتين $n$ و $m$ ، وتكون مقبولة إلى حدٍّ ما عندما تكون قيمة كلّ من $n$ و $m$ أكبر من 32.

٢ – إذا كان الطرف الأيمن من الفترة الأخيرة [10,34] سالباً فإنَّه سيكون لدينا $p_1<p_2$ باحتمال قدره $(1-\alpha)$ .

٣ – إذا كان الطرف الأيسر من الفترة الأخيرة [10,34] موجباً فإنَّه سيكون لدينا $p_1>p_2$ باحتمال قدره $(1-\alpha)$ .

٤ – إذا كان الطرف الأيمن للفترة الأخيرة [10,34] موجباً وطرفها الأيسر سالباً، فعندئذ يكون $p_1=p_2$ باحتمال قدره $(1-\alpha)$ ، ومن ثمّ فإنَّ ما يظهر من فوارق في قيم الإحصاءتين $\overline{\pi}_1$ و $\overline{\pi}_2$ هي فوارق ظاهريّة لا أهمّيّة لها ما دامت قيمة $(1-\alpha)$ قريبة من الواحد.

**مثال (١٠,٣,٥,٣)**

في مصنع للأدوات الكهربائية توجد آلتان $M_1$ و $M_2$ لإنتاج الملفات المستخدمة في تصنيع المحركات الكهربائية، وتعمل كلّ منها بشكل مستقلّ عن الأخرى، قمنا بسحب عيِّنة من إنتاج $M_1$ بحجم $n = 196$ ملفاً فوجدنا فيها 13 ملفاً غير صالحة للعمل، وكذلك قمنا بسحب عيِّنة أخرى من إنتاج $M_2$ بحجم $m = 225$ ملفاً فوجدنا فيها 15 ملفاً غير صالحة للعمل. هل تدلّ هذه المعطيات على أنَّ للآلتين $M_1$ و $M_2$ الكفاءة نفسها في الإنتاج عند مستوى الثّقة 98%؟

**الحل:** من أجل ذلك لنقم بتعيين 98% فترة ثقة للفرق $p_1 - p_2$ علماً أنَّ الفترة المطلوبة لذلك تُعطى بالفترة [10,34] حيث لدينا من مُعطيات المسألة $\bar{p}_1 = \frac{13}{196} = 0.066$ و $\bar{p}_2 = \frac{15}{225} = 0.067$، ومن جدول التوزيع الطبيعي المعيّاري نجد أنَّ $z_{1-\frac{\alpha}{2}} = z_{0.99} = 2.33$، ومن ثَمَّ نجد أنَّ لحدّي الفترة [10,34] العرض الآتي:

$$(0.066 - 0.067) \mp 2.33\sqrt{\frac{0.066(1-0.066)}{196} + \frac{0.067(1-0.067)}{225}}$$

والتي نجد منها أنَّ 98% فترة ثقة للفرق $p_1 - p_2$ هي $(-0.0556\ ,\ 0.0576)$، وهذا يعني أنَّه للآلتين $M_1$ و $M_2$ الكفاءة نفسها في الإنتاج عند مستوى الثّقة 98%؛ وذلك لأنَّ الحد الأدنى لهذه الفترة سالب في حين حدّها الأعلى موجب.

نقدِّم فيما يلي بعض التطبيقات على استخدام برنامج Minitab في مجال تعيين فترات الثّقة.

## (١٠,٤) تطبيقات

**(١٠,٤,١) تطبيق**

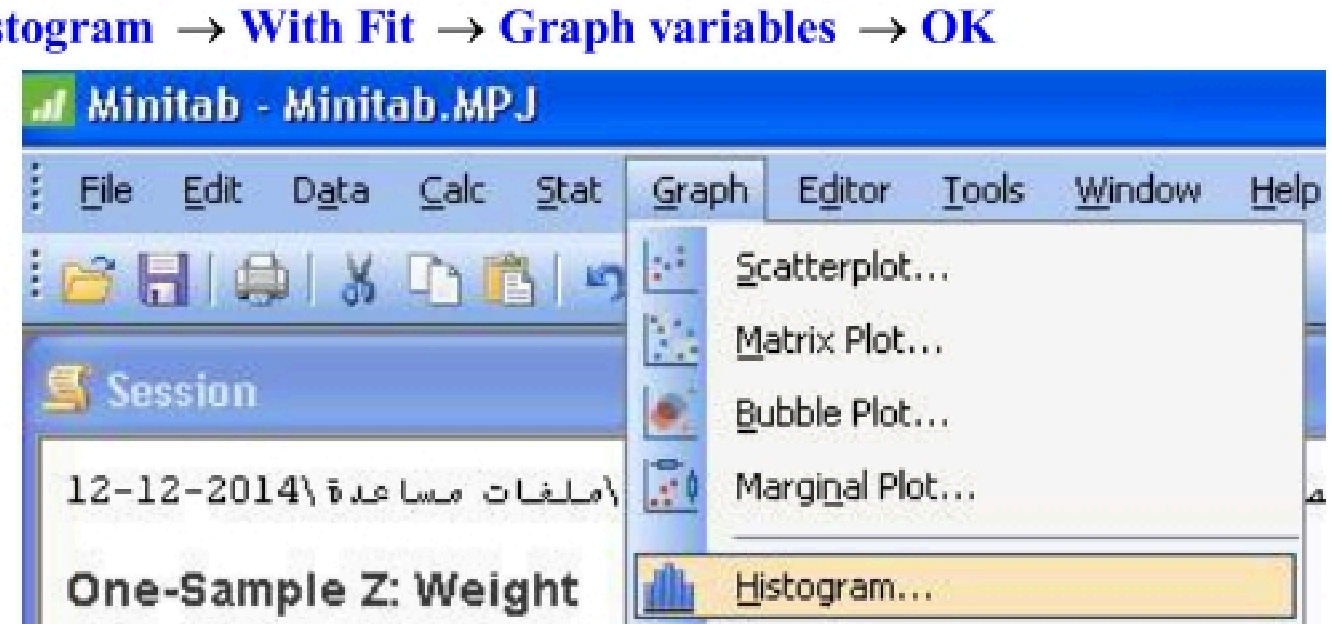
لتكن لدينا البيانات الآتية التي تمثّل الأوزان لـ 64 شخصاً بالغاً مقدَّرة بالكيلو غرام:

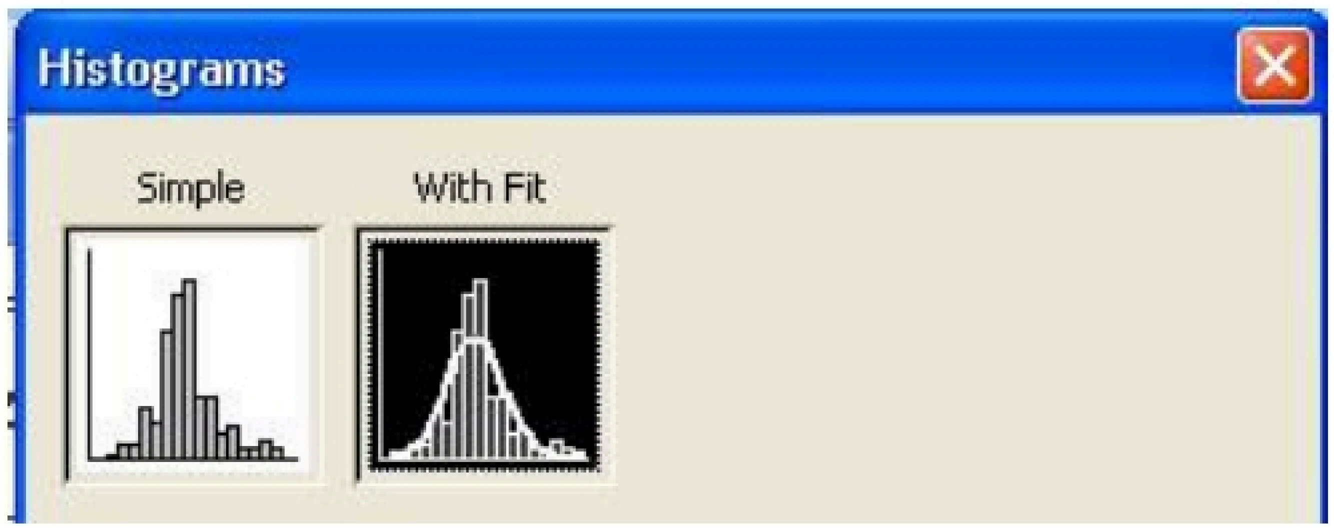
| | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 69 | 94 | 83 | 59 | 79 | 74 | 73 | 87 | 84 | 90 | 89 | 81 | 70 | 53 | 71 | 61 |
| 76 | 87 | 65 | 61 | 78 | 83 | 75 | 88 | 65 | 93 | 66 | 66 | 99 | 57 | 98 | 60 |
| 59 | 72 | 80 | 58 | 91 | 68 | 82 | 67 | 79 | 52 | 74 | 76 | 77 | 73 | 80 | 89 |
| 78 | 72 | 62 | 96 | 73 | 82 | 55 | 56 | 86 | 84 | 54 | 63 | 85 | 76 | 76 | 64 |

ولننظر أولاً إن كانت هذه البيانات تتوزَّع طبيعياً أو لا.

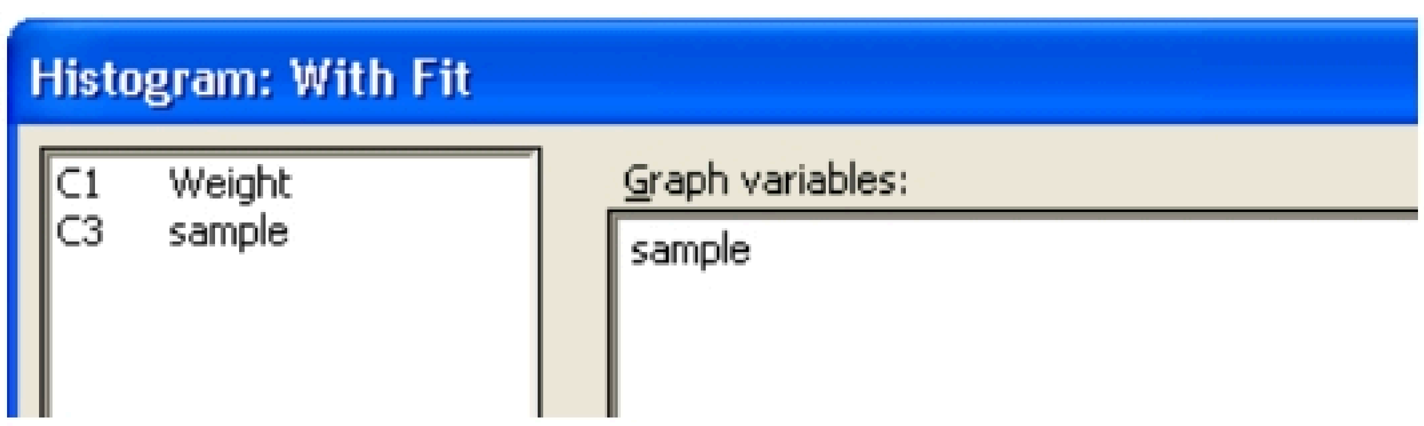
من أجل ذلك سنحاول الاستدلال على شكل التوزيع من خلال مدرَّجه التكراري ومقارنته مع التوزيع الطبيعي مستخدمين في ذلك برنامج Minitab وفقاً للخطوات الآتية:

**Graph → Histogram → With Fit → Graph variables → OK**

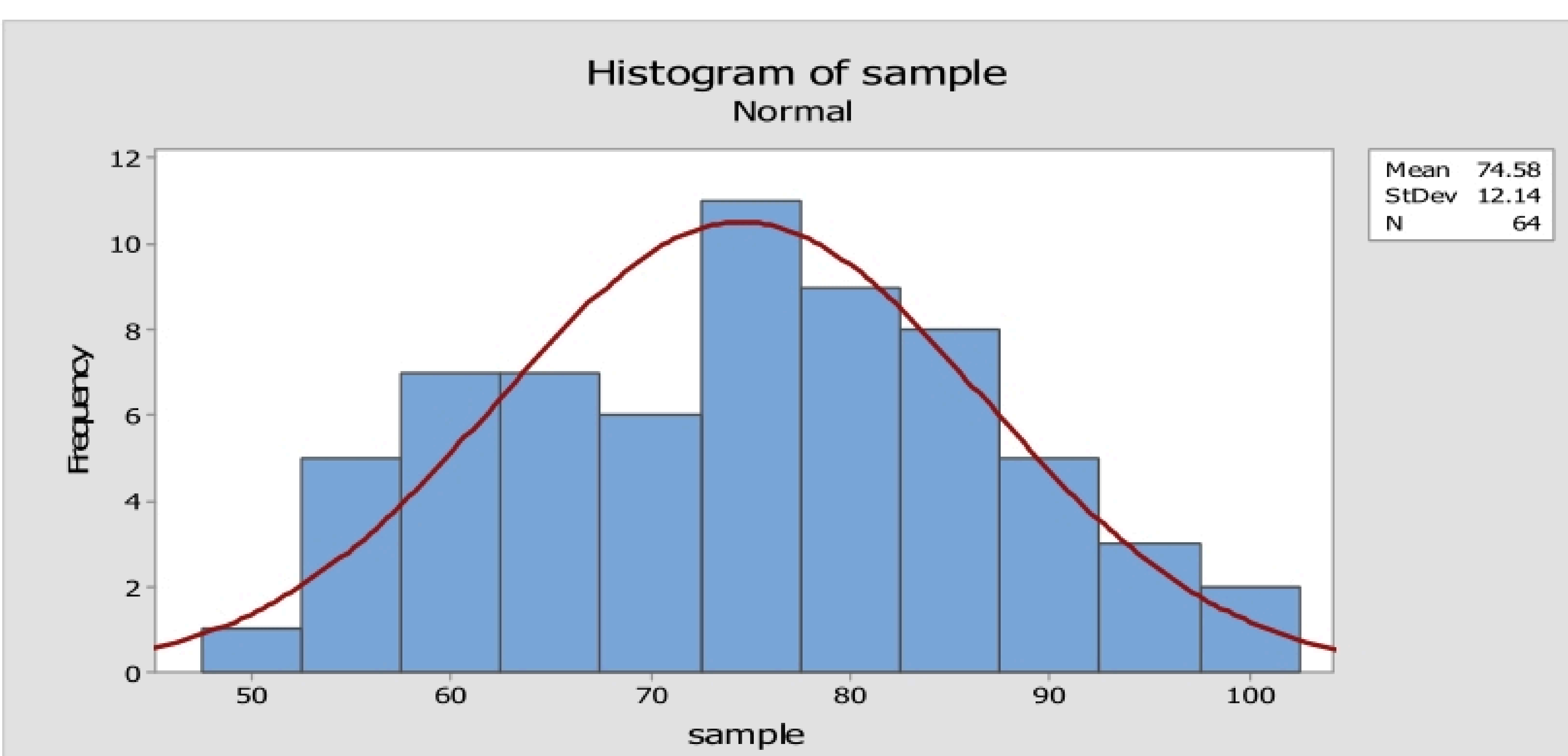





ومن ثمَّ




وبعد ذلك تفعيل OK فنحصل على الشكل الآتي المُمثِّل للمدرَّج التكراري مع منحنى التوزيع الطبيعي الملائم لهذه البيانات:

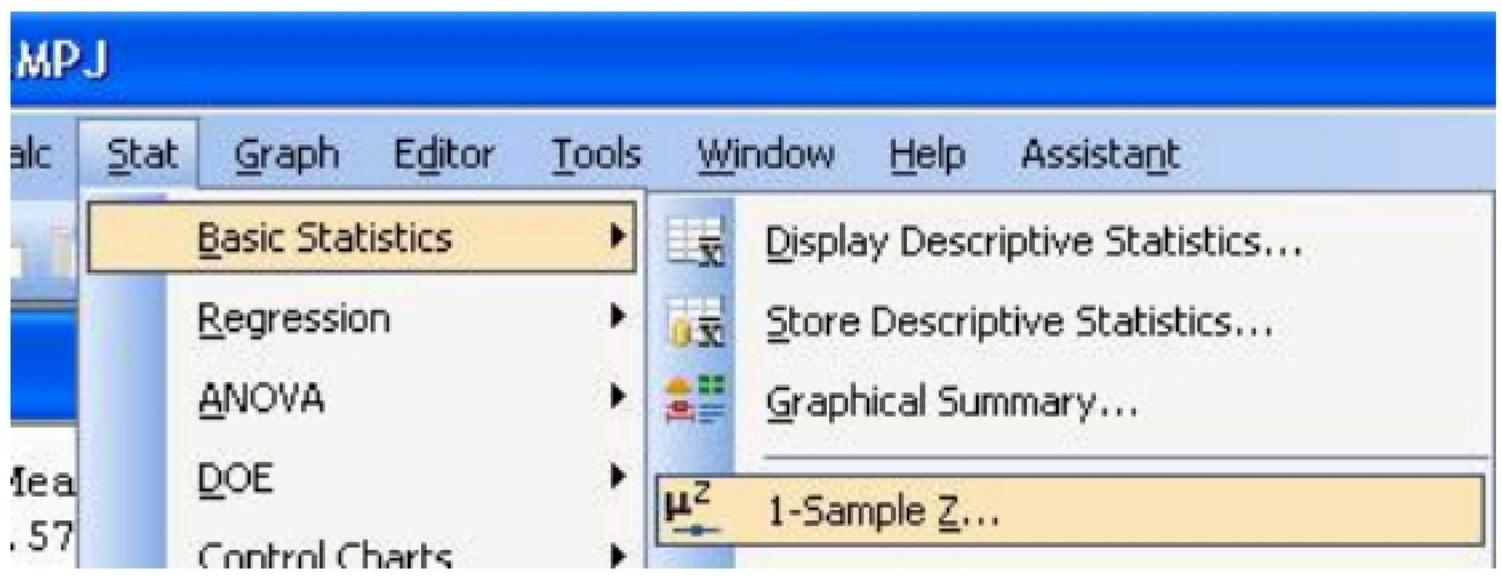



فنلاحظ أنَّ هذا العرض يظهر تقارباً جيداً مع منحنى دالَّة الكثافة الاحتمالية للتوزيع الطبيعي، ويزودنا هذا العرض بقيمة المتوسط $\bar{x} = 74.578$ والانحراف المعياري $s = 12.135$، وأمّا من أجل اختبار إن كانت هذه البيانات تقبل التوزيع الطبيعي كتوزيع مناسب فعلاً فإنَّنا سنترك ذلك إلى التطبيقات المقدَّمة في نهاية الفصل القادم (اختبار الفرضيات الإحصائية).

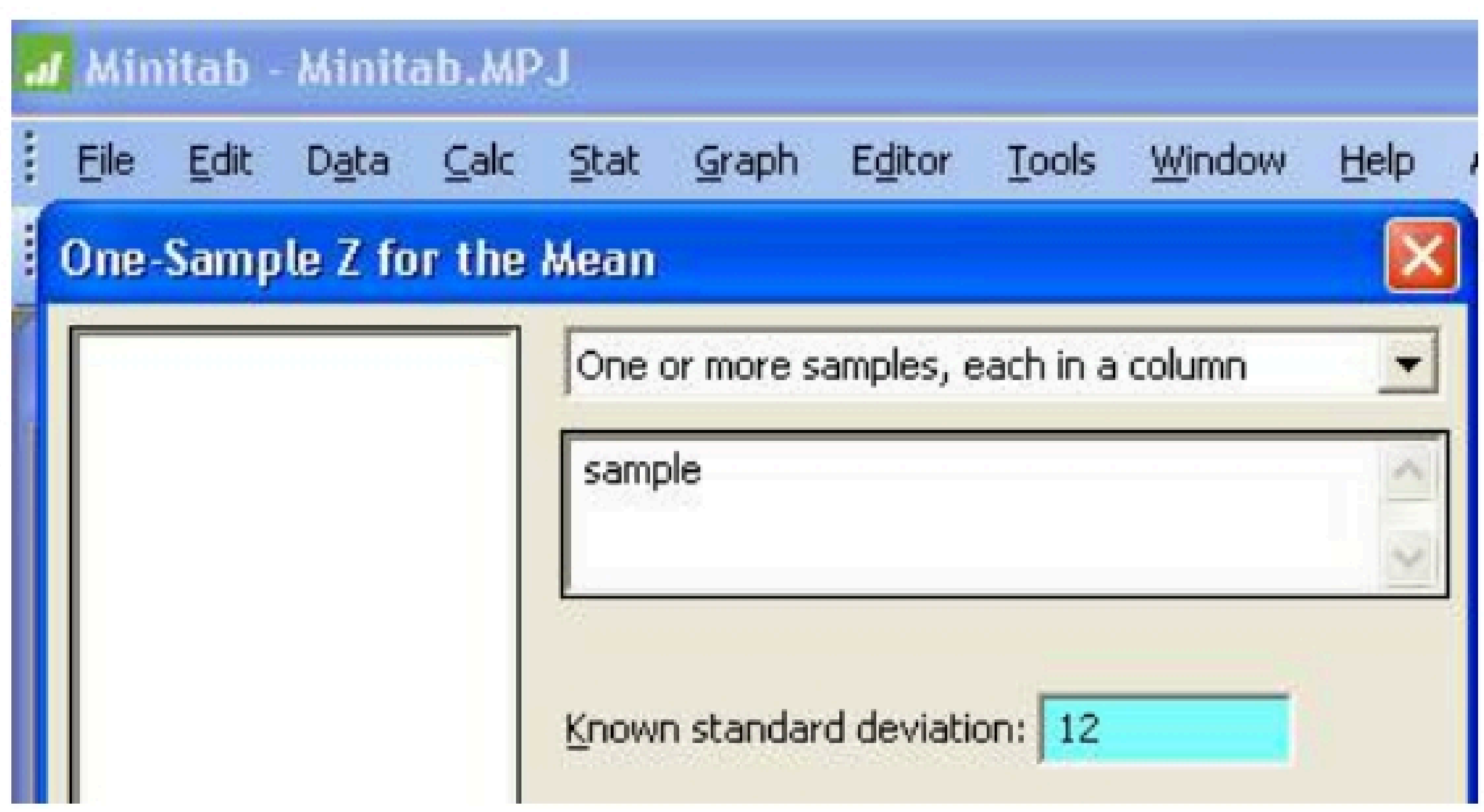
الآن لتعيين $\%(1-\alpha)100$ فترة ثقة لـ $\mu$ متوسط مجتمع هذه البيانات، فإنَّنا نعلم أنَّه علينا أن نميِّز بين حالتين هما:

أ- إذا افترضنا أنَّ الانحراف المعياري للمجتمع معلوم ويساوي $\sigma = \sigma_o = 12$، فعندئذ نتبع الخطوات الآتية (مستخدمين توزيع $Z$ - أو الطبيعي-):

**Stat → Basic Statistics → 1-Sample Z → Known standard deviation → OK.**




ومن ثمَّ الخطوة التالية



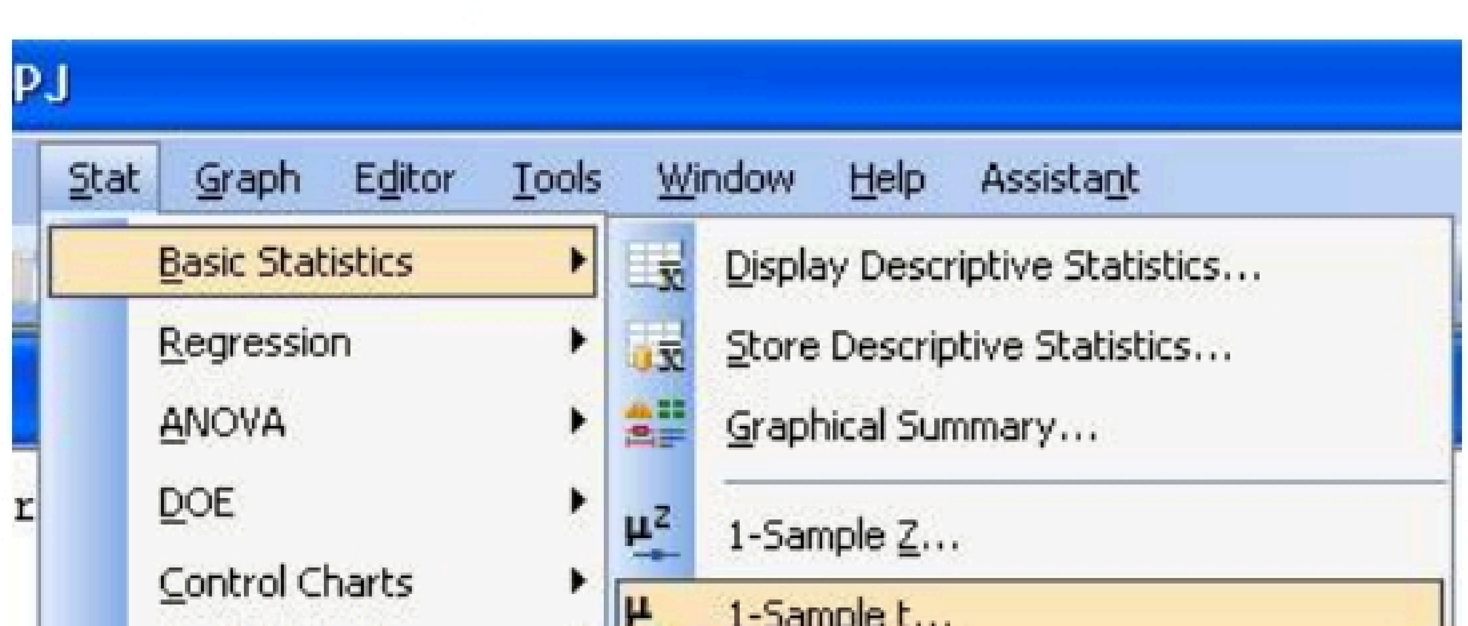

وبعد ذلك تفعيل OK فنحصل على التقرير الآتي الذي يحتوي على 95% فترة ثقة للمتوسط $\mu$ .

**One-Sample Z: sample**
**The assumed standard deviation = 12**

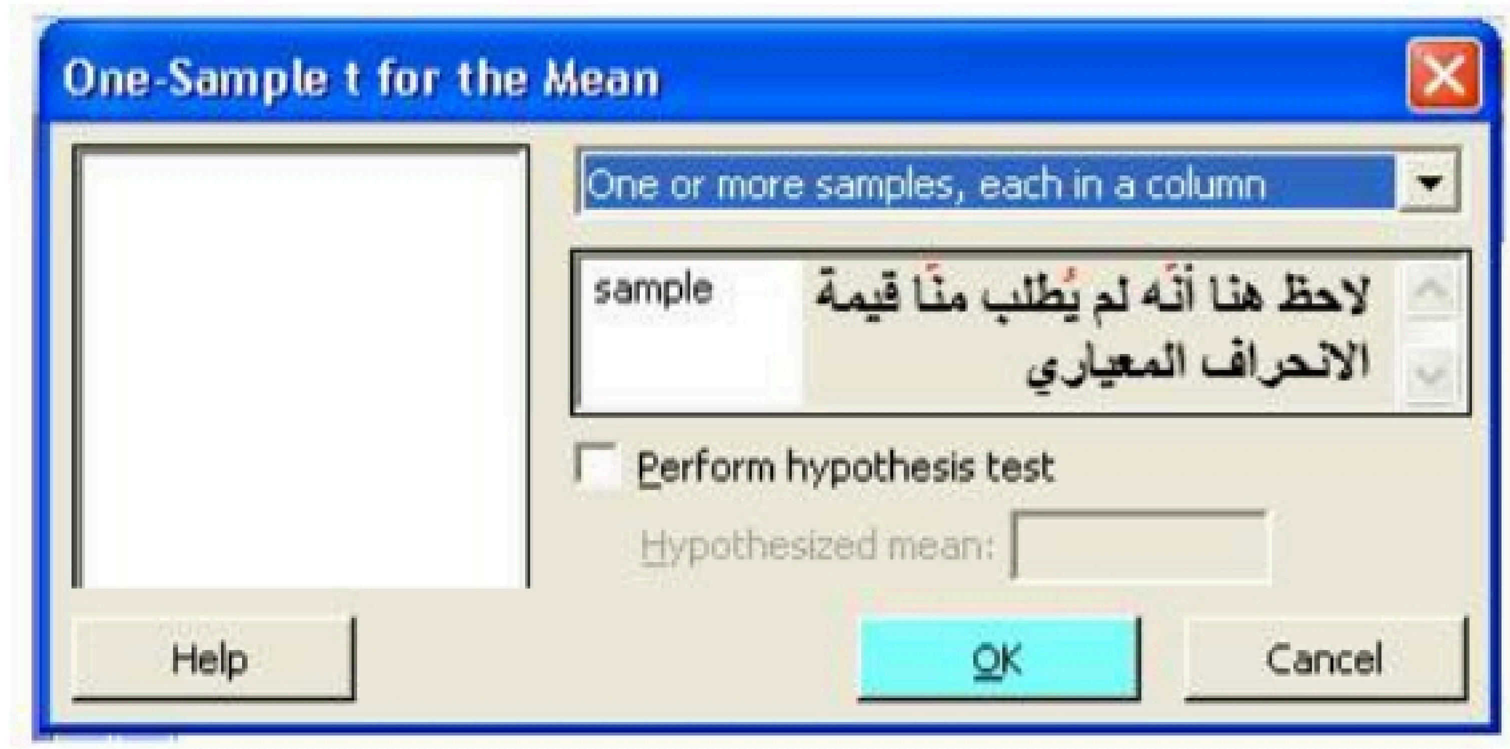
| Variable | N | Mean | St Dev | SE Mean | 95% CI |
|---|---|---|---|---|---|
| sample | 64 | 74.58 | 12.14 | 1.50 | (71.64 , 77.52) |

ب- إذا افترضنا أنَّ الانحراف المعياري للمجتمع مجهول $\sigma = ?$ ، فعندئذ نتبع الخطوات الآتية (مستخدمين توزيع **t** – أو ستودنت-):

**Stat → Basic Statistics → 1-Sample t → Known standard deviation → OK**




ومن ثمَّ الخطوة التالية



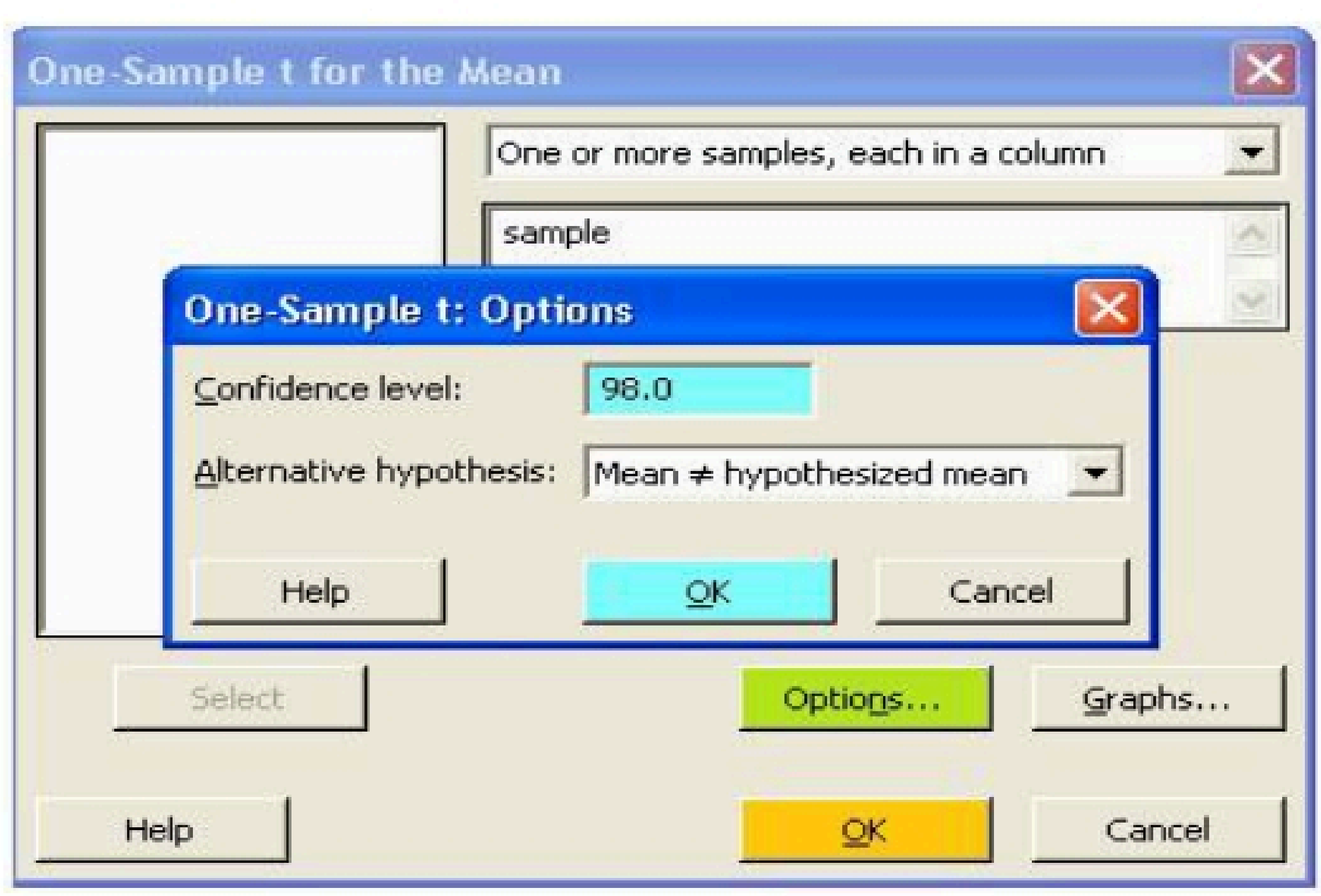

وبتفعيل OK نحصل على التقرير الآتي الذي يحتوي على 95% فترة ثقة للمتوسط $\mu$ .

**One-Sample T: sample**

| Variable | N | Mean | St Dev | SE Mean | 95% CI |
|---|---|---|---|---|---|
| sample | 64 | 74.58 | 12.14 | 1.52 | (71.55 , 77.61) |

وإذا كنّا نرغب في تغير مستوى الثّقة فإنّه يمكننا التعديل على هذه القيمة من خلال المرحلة قبل الأخيرة على النحو الآتي:

**One - Sample t for the mean → Option → Confidence level → OK → OK**




فنحصل على التقرير الآتي الذي يحتوي على 98% فترة ثقة للمتوسط $\mu$ حسب كل حالة على حدة.

**One-Sample Z: sample, The assumed standard deviation = 12**

| Variable | N | Mean | St Dev | SE Mean | 98% CI |
|---|---|---|---|---|---|
| sample | 64 | 74.58 | 12.14 | 1.50 | (71.09 , 78.07) |

**One-Sample T: sample, The assumed standard deviation = 12**

| Variable | N | Mean | St Dev | SE Mean | 98% CI |
|---|---|---|---|---|---|
| sample | 64 | 74.58 | 12.14 | 1.52 | (70.96 , 78.20) |

وأمّا لتعيين 95% فترة ثقة لـ $\sigma^2$ تباين مجتمع البيانات المقدّمة فإنّنا نتبع الخطوات الآتية:

**Stat → Basic Statistics → 1Variance → Insert sample in free space → OK.**

فنحصل على التقرير الآتي الذي يحتوي على 95% فترة ثقة لكل من $\sigma$ الانحراف المعياري و $\sigma^2$ التباين لمجتمع البيانات المقدّمة.

**Test and CI for One Variance: sample**
**Method: The chi-square method is only for the normal distribution.**

| **Variable** | **N** | **St Dev** | **Variance** |
|---|---|---|---|
| **sample** | **64** | **12.1** | **147** |

| **95% Confidence Intervals** | **St Dev** | **Variance** |
|---|---|---|
| | **(10.3; 14.7)** | **(106.09; 216.09)** |

**(١,١,٤,١٠) ملاحظة**

بأسلوب مماثل يمكننا الحصول على العروض وفترات الثّقة الخاصّة بفرق المتوسطين $\mu_1-\mu_2$ ونسبة التباين $\sigma_1^2/\sigma_2^2$ لمجتمعين طبيعيين، ولكن يجب الانتباه إلى أنَّ هذا البرنامج يشترط أن يكون للعينتين اللتين ستخضعان للدراسة الحجم نفسه.

**(١٠,٤,٢) تطبيق**

لتكن لدينا البيانات الآتية ( **تحت مُسمَّى Bi-sample** ) التي تمثِّل عيِّنة بحجم 81 شخصاً استطلعت آراءهم حول مشروع تنموي، علماً أنَّ الإجابة "**موافق**" مُثِّلت بالعدد 1 في حين مُثِّلت الإجابة "**غير موافق**" بالعدد 0.

| | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | |
| 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | |
| 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 0 | 0 | 1 | |
| 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

فعندئذ يمكننا استخدام برنامج Minitab للحصول على $(1-\alpha)100\%$ فترة ثقة لـ $p$ مَعْلمة مجتمع هذه البيانات وذلك وفقاً للخطوات الآتية:

**Stat → Basic Statistics → 1-Proportion → Insert Bi-sample in free space → OK.**

فنحصل على التقرير الآتي الذي يحتوي على 95% فترة ثقة لـ $p$ مَعْلمة مجتمع هذه البيانات المقدَّمة.

**Test and CI for One Proportion: Bi-sample**
**Event = 1**

| **Variable** | **X** | **N** | **Sample p** | **95% CI** |
|---|---|---|---|---|
| **Bi-sample** | **68** | **81** | **0.839506** | **(0.741200; 0.911681)** |

**(١,٢,٤,١٠) ملاحظة**

بأسلوب مماثل تماماً يمكننا الحصول على العروض وفترات الثّقة الخاصّة بفرق الوسطين $p_1-p_2$ لمجتمعين برنوليين، ولكن يجب الانتباه هنا إلى أنَّ هذا البرنامج يشترط أن يكون للعينتين اللتين ستخضعان للدراسة الحجم نفسه أيضاً.

## هذا ما تيسَّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل

# تمارين الفصل العاشر

١- ليكن $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع التوزيع المنتظم المستمر على فترة $[a, b]$ مع $-\infty < a < b < +\infty$، ولنأخذ $\mathcal{X}_n$ عيِّنة من هذا المجتمع، والمطلوب استخدام طريقة العزوم في تعيين مُقدِّرات مَعْلمتي هذا المجتمع.

٢- ليكن $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ بدالَّة كثافة احتمالية $f_X(x\,;\theta)$ مُعطاة من خلال العلاقة:

$$f_X(x\,;\theta) = \frac{1}{\theta}\mathbf{I}_{[0,\theta]}(x)$$

ولتكن $\mathcal{X}_n$ عيِّنة من هذا المجتمع، والمطلوب تعيين مُقدِّر الأرجحية العظمى للمَعْلمة $\theta$.

٣- ليكن $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولنأخذ $\mathcal{X}_n$ عيِّنة من هذا المجتمع، والمطلوب استخدام طريقة العزوم في تعيين مُقدِّرات مَعْلمة هذا المجتمع في الحالتين الآتيتين:

أ- المجتمع يتبع التوزيع المنتظم المتقطع بمَعْلمة $n \in \mathbb{N}$.

ب- المجتمع يتبع التوزيع الهندسي بمَعْلمة $0 < p < 1$.

ومن ثمَّ بين أيّ من هذه المقدِّرات هو مُقدِّر غير منحاز، وأيُّها غير منحاز تقاربيًّا.

٤- ليكن $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولنأخذ $\mathcal{X}_n$ عيِّنة من هذا المجتمع، فعندئذ استخدم طريقة الأرجحية العظمى في تعيين مُقدِّرات مَعْلمة هذا المجتمع في الحالات الآتية:

أ- المجتمع يتبع التوزيع الأسي بمَعْلمة $\lambda > 0$.

ب- المجتمع يتبع توزيع برنولي بمَعْلمة $0 < p < 1$.

ج- المجتمع يتبع توزيع له دالَّة كثافة احتمالية:

$$f_X(x\,;\theta) = \theta\, x^{\theta-1}\,\mathbf{I}_{(0,1)}(x) \quad ;\theta > 0 \;\&\; x \in \mathbb{R}$$

د- المجتمع يتبع توزيع له دالَّة كثافة احتمالية:

$$f_X(x\,;\theta) = \theta\,(1-x)^{\theta-1}\,\mathbf{I}_{(0,1)}(x) \quad ;\theta > 0 \;\&\; x \in \mathbb{R}$$

هـ- المجتمع يتبع توزيع له دالَّة كثافة احتمالية:

$$f_X(x\,;\theta) = \mathbf{e}^{\theta-x}\;\mathbf{I}_{(\theta,\infty)}(x) \quad ;\theta, x \in \mathbb{R}$$

ومن ثمَّ بين أيّ من هذه المقدِّرات هو مُقدِّر غير منحاز، وأيُّها غير منحاز تقاربيًّا.

٥- زعم أحد المتاجر أنَّ متوسط أرباحه اليومية على مرور أيام السنة يساوي 1000 ريال سعودي، وبانحراف معياري قدره 25 ريالًا سعودي وذلك عند مستوى من الثِّقة قدره 98%. الآن، وللتأكد من صحة زعمه قمنا بحساب الأرباح اليومية لستة عشر يوماً متفرقة من السنة، فوجدنا المُعطيات الآتية:

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1000 | 1000 | 1010 | 900 | 1010 | 1100 | 1000 | 1000 |
| 800 | 950 | 850 | 950 | 950 | 1200 | 800 | 1300 |

والمطلوب ما يلي:

أ- إذا أخذنا بالمعطيات التي قدَّمها صاحب المتجر، فهل يمكن القبول بزعمه حول متوسط أرباحه اليومية عند مستوى من الثّقة قدره 98%؟

ب- إذا تجاهلنا المُعطيات التي قدَّمها صاحب المتجر، فهل يمكن القبول بزعمه حول متوسط أرباحه اليومية عند مستوى من الثّقة قدره 98%؟

ج- ما هو الخطأ المرتكب في تقدير متوسط الأرباح اليومية للبائع بقيمة الإحصاءة $\overline{X}$ في كلٍّ من الحالتين الآتيتين:

- الأخذ بقيمة الانحراف المعياري المقدَّمة من قبل صاحب المتجر.
- تجاهل قيمة الانحراف المعياري المقدَّمة من قبل صاحب المتجر.

٦- في مؤسسة صناعية يوجد الآلاف من العاملين، فإذا علمنا أنَّ الأجر الشهري للعمال يخضع للتوزيع الطبيعي $\mathbf{N}(\mu, 30)$، فعندئذ كم يجب أن يكون حجم العيِّنة المسحوبة من مجتمع العاملين بحيث إنه عند مستوى ثقة قدره 98% يكون الخطأ المرتكب في تقديرنا لقيمة $\mu$ بالقيمة $\overline{x} = 650$ يساوي $\pm 25$ وحدة نقدية؟

٧- ادعى مصنع لإنتاج المدخرات الكهربائية ذات الجهد 12.3 V أنَّ متوسط عمر المُدَّخرات لديه يساوي 750 ساعة عمل بانحراف معياري قدره 13 ساعة عمل. الآن إذا افترضنا أنَّه يوجد 1000 مُدَّخرة في مستودع إنتاج هذا المصنع، وقمنا بأخذ عيِّنة من هذه المدخرات بحجم 36 مُدَّخرة واستخدامها، فوجدنا أنَّ متوسط عمر هذه المُدَّخرات يساوي 725 ساعة عمل بانحراف معياري قدره 15 ساعة عمل، فإذا كانت أعمار المدخرات تخضع للتوزيع الطبيعي بمتوسط $\mu$، وانحراف معياري $\sigma$، فعندئذ عيِّن 90% فترة ثقة لمتوسط أعمار المدخرات الكهربائية في كلٍّ من الحالات الاتية:

أ- إذا تجاهلنا قيمة الانحراف المعياري المقدَّمة من قبل المصنع.

ب- إذا أخذنا بقيمة الانحراف المعياري المقدَّمة من قبل المصنع.

٨- يَدَّعي مصنع لإنتاج المكثِّفات الكيميائية (تستخدم في الدوائر الإلكترونية والكهربائية) ذات السعة $1000\mu F$ أنَّ نسبة القطع المعطَّلة في إنتاجه لا يتجاوز 0.025، وذلك عند مستوى من الثّقة مقداره 95%، فإذا افترضنا أنَّ لجميع المكثفات النصيب نفسه في السحب، وقمنا بسحب عيِّنة من إنتاج هذا المصنع بحجم 225 مكثِّفاً فوجدنا لدى فحصها 13 مكثِّفاً معيباً، والمطلوب:

أ- تحقَّق من صحة ادعاء المصنع عند مستوى من الثّقة مقداره 95%.

ب- ما هو حجم العيّنة التي يجب سحبها بحيث لا يتجاوز الخطأ المرتكب في تقديرنا لَمْلمة مجتمع المكثِّفات $p$ بالقيمة $\overline{p} = 0.001$ (التي تمثِّل التكرار النسبي للمكثِّفات المعطَّلة التي ستظهر لدينا في العيّنة) بأكثر من $\pm 0.01$؟

٩- نريد القيام بالتّحقُّق من نتيجة تجربة ما بحيث إنه عند مستوى من الثّقة 90 % تكون الفوارق التي سنحصل عليها في حدود $\pm 0.06$ من النسبة الحقيقية لنجاح هذه التجربة، فإذا كانت التجارب الخاصة بهذه المسألة تُكرَّر بشكل مستقل بعضها عن بعض، وتحت الشروط نفسها، فعندئذ كم مرَّة يجب أنَّ نُكرِّر التجربة عند مستوى ثقة قدره 90 % لكي يكون الفارق في حدود $\pm 0.08$ من النسبة الحقيقية لنجاح هذه التجربة:

أ- إذا علمنا من باحثين آخرين أنَّ نسبة النجاح لهذه التجربة هو 0.30؟

ب- إذا كنّا لا نعلم شيئاً عن نسبة النجاح لهذه التجربة؟

١٠- عُرض على إحدى شركات تجميع الحواسيب نوعاً من السوّاقات الليزرية من مصدرين $A$ و$B$ مستقلّين بعضهما عن بعض، حيث ادّعى المصدر $A$ أنّه على ثقة قدرها 0.99 بأنّ متوسط عمر السواقة لديه 1000 ساعة عمل بانحراف معياري قدره 15 ساعة عمل، بينما ادعى المصدر $B$ أنّه على ثقة قدرها 0.99 بأنّ متوسط عمر السواقة لديه 1050 ساعة عمل بانحراف معياري قدره 5 ساعات عمل. فلو افترضنا أنّ عمر السوّاقات يخضع للتوزيع الطبيعي، فأي المصدرين يُفضَّل، وذلك عند مستوى من الثّقة قدره 0.99؟

١١- يقوم باحثان $A$ و $B$ كل منهما بشكل مستقل عن الآخر، بإجراء تجربة معينة وتحت الشروط نفسها، فقد قام الباحث $A$ بتكرارات مستقلّة لهذه التجربة عددها $n$ مرّة حيث تحقّقت النتائج المرجوة منها في $n_1$ مرّة، وأمّا الباحث الآخر $B$ فقد قام بتكرارات مستقلّة لهذه التجربة عددها $m$ مرّة حيث تحقّقت النتائج المرجوة منها في $m_1$ مرّة. عندئذ المطلوب ما يلي:

أ- إذا كان $n = m = 49$ و $n_1 = m_1 = 36$، فعندئذ عيّن 98% فترة ثقة للفرق بين وسيطي النجاح للتجربة التي نُفّذت من قبل الباحثين.

ب- إذا كان $n = 81$ و $n_1 = 64$ و $m = 100$ و $m_1 = 49$، فعندئذ:

- بيّن عند أي مستوى من الثّقة يمكننا أن نؤكد على أنّ قيمة مَعْلمة نجاح التجارب لدى الباحث الأول $A$ أصغر من قيمة مَعْلمة النجاح لتجارب الباحث الآخر $B$.
- عند أي مستوى من الثّقة يكون الخطأ المرتكب في تقديرنا للفرق $p_1 - p_2$ بقيمة الإحصائي $\bar{P}_1 - \bar{P}_2$ لا يتجاوز $\pm 0.005$.

١٢- في مصنع للحديد والصلب لدينا خطين مستقلّين بعضهما عن الآخر لإنتاج القضبان الحديدية ذات الطول 5 أمتار وبقطر يساوي 12 مم. قمنا بسحب عيّنة $\mathfrak{X}_n$ من إنتاج الخط الأول وعيّنة عشوائية أخرى $\mathfrak{Y}_m$ من إنتاج الخط الثاني، فوجد أنّ متوسط قطر القضبان في العيّنة الأولى 12.1 مم بانحراف معياري قدره 0.5 مم وأمّا متوسط قطر القضبان في العيّنة الثانية فكان 11.9 مم بانحراف معياري قدره 0.2 مم، والمطلوب ما يلي:

أ- إذا كان $n = 49$ و $m = 36$، فعندئذ عين 95% فترة ثقة لفرق متوسط قطر القضبان الناتجة من الخطين إذا أخذنا بقيم الانحراف المعياري المقدَّمة من قبل المصنع، ومرّة أخرى إذا تجاهلنا قيم الانحراف المقدّمة من قبل المصنع.

ج- هل تتوافق النتائج التي حصلنا عليها من العينات حول الانحراف المعياري مع ما قدّمه المصدرين $A$ و$B$ وذلك عند مستوى من الثّقة قدره 95%؟

د- إذا كان $n = 49$ و $m = 36$، فعند أي مستوى من الثّقة يمكننا البت في أنّ متوسط قطر القضبان الناتج عن الخط الأول أكبر من متوسط قطر القضبان الناتج عن الخط الثاني؟

# الفصل الحادي عشر

## مدخل إلى اختبارات الفرضيات الإحصائية

## AN INTRODUCTION TO STATISTICAL HYPOTHESES TESTES

### (١١,١) مفاهيم أساس لدراسة اختبار فرضية إحصائية

### Based Concepts for Study a Statistical Hypothesis Test

لقد قدَّمنا في الفصل السابق فترات الثقة لمَعَالم مجتمع (طبيعي أو برنولي)، وقد لاحظنا أنَّ هذه الفترات قد تساعدنا في بعض الحالات على إصدار قرار حول مقولة متعلِّقة بالمجتمع (وللتذكير نقصد بمجتمعاً إحصائيًاً دوماً). لكن الهدف الأساس من تلك الفترات كان تعيين الفترة التي تنتمي إليها مَعْلَمة المجتمع باحتمال معلوم، ومن ثمَّ تنفيذ بعض العمليّات المتعلِّقة بذلك مثل تعيين الخطأ الأعظمي المرتكب نتيجة أخذ قيمة إحصاءة مقدِّرة لمَعْلَمة كقيمة بديلة لهذه المَعْلَمة، وكذلك تعيين حجم العيِّنة الذي يجب أن يستخدم بحيث لا يتجاوز الخطأ مقداراً مُحدَّداً مسبقاً. أمَّا مسألة اتخاذ قرار حول قبول أو رفض مقولة متعلِّقة بالمجتمع المدروس (كأن تكون متعلِّقة بخاصيةٍ ما من خصائصه مثل: المتوسط، أو الانحراف المعياري، أو التوزيع، أو...) مع الأخذ بالحسبان مقدار احتمال رفضها وهي صحيحة، أو قبولها وهي خاطئة فإنَّ ذلك يخرج عن إمكانيات فترات الثقة، والقطاع الإحصائي الذي يهتمّ بمثل هذه المسائل يُعرف باسم اختبار الفرضيّات الإحصائية.

في الواقع يُنظر إلى اختبارات الفرضيّات الإحصائية على أنَّه أحد جوانب الاستدلال الإحصائي وركيزة أساس لنظرية اتخاذ القرار Decision-Making Theory، علماً أنَّ الاستدلال الإحصائي ونظرية اتخاذ القرار يهتمان بشكل أساس بالطرائق والتطبيقات المكوِّنة للأحكام حول خصائص المجتمع معتمدة في ذلك على عيِّنات (وللتذكيرنقصد بها عيِّنات عشوائيّة دوماً) مسحوبة من ذلك المجتمع. أكثر من ذلك، فإنَّ اختبارات الفرضيّات الإحصائية تهتمّ بإنجاز تلك العمليّات التي تقوم على قبول أو رفض فرضيّة أو مقولة متعلِّقة بالمجتمع الذي قيد الدراسة معتمدة في ذلك على عيِّنات مسحوبة من هذا المجتمع، فإذا كانت الفرضيّة التي تحت الاختبار تتعلَّق بمَعْلَمة أو أكثر من مَعَالم المجتمع المدروس، فعندئذ يُدعى هذا الاختبار بـ اختبار مَعْلمي Parametric Test، وأمَّا إذا كانت الفرضيّة التي تحت الاختبار تتعلَّق بتوزيع المجتمع أو بخاصيّة ما مثل الاستقلال بين التوزيعات أو التّجانس، فعندئذ يُدعى مثل هذا الاختبار بـ اختبار غير مَعْلمي Non Parametric Test.

في هذا الفصل سنقدِّم بعض الاختبارات المَعْلميّة وأخرى غير المَعْلميّة، وهذه الاختبارات تتطلب منّا تقديم بعض المفاهيم المتعلِّقة باختبار الفرضيّات، حيث سنستند في معلوماتنا المقدَّمة في هذا الفصل على المراجع الآتية من المراجع المذكورة في نهاية هذا الكتاب [6]، [20]، [24]، [33]، [43]، [45]، [47]، [55]، [62] و[63].

#### (١١,١,١) الفرضيّات الإحصائية Statistical Hypotheses

إنَّ أول ما يتبادر لذهن المرء هو معرفة ما تعنيه عبارة فرضيّة إحصائية، ولذلك سنبدأ بتقديم مجموعة من المفاهيم التي تخدم هذا المصطلح.

**(١١,١,١,١) تعريف (الفرضيّة الإحصائية)**

إنَّ **الفرضيّة الإحصائية** هي مقولة أو إفادة تتعلّق بالمجتمع المدروس، وهذه المقولة تحتمل الصحة أو الخطأ.

**(١١,١,١,٢) ملاحظات**

١- تجدر الإ شارة هنا إلى أنَّ قبول الفرضيّة الإحصائية لا يعني بالضرورة أنَّها صحيحة، ولكن وفقاً لمعطيات العيّنة المسحوبة من المجتمع ليس لدينا ما يدفعنا إلى رفضها، في حين أنَّ رفض الفرضيّة الإحصائية إنَّما يعني الإقرار بخطئها حتماً.

٢- عادةً يضع المرء الفرضيّة التي يأمل أن تكون صحيحة مَوْضعَ الرفض ويُطلق عليها اسم **الفرضيّة الابتدائية** Null Hypothesis (أو الفرضيّة الصفرية-أو فرضيّة العدم) ويُرْمَز لها بـ $\mathbf{H}_o$، وإذا ما توصَّل المرء (نتيجة لمُعطيات العيّنة المسحوبة من المجتمع) إلى رفض هذه الفرضيّة فإنَّه سيلجأ إلى قبول فرضيّة أخرى يُطلق عليها اسم **الفرضيّة البديلة** Alternative Hypothesis ويُرْمَز لها عادة بـ $\mathbf{H}_A$، والأمثلة الآتية توضِّح لنا ذلك.

**(١١,١,١,٣) أمثلة**

١- لنأخذ تجربة إلقاء حجر نرد لمَرَّة واحدة فقط، فعندئذ بفرض أنَّ $p$ هو احتمال ظهور أيّة قيمة من القيم الست الممكنة للأعلى، فعندئذ بوضع $\mathbf{H}_o : p = \frac{1}{6}$، فإنَّ ذلك يعني أنَّه لدينا اختبار للفرضيّة الابتدائية القائلة بتوازن حجر النرد، وفي حال التوصل إلى رفض هذه الفرضيّة (نتيجة لمُعطيات العيّنة) فإنَّه يجب القبول بفرضية أخرى بديلة $\mathbf{H}_A : p \neq \frac{1}{6}$ أو $\mathbf{H}_A : p > \frac{1}{6}$ أو $\mathbf{H}_A : p < \frac{1}{6}$ وكلّ منها تُعبِّر عن عدم توازن حجر نرد.

٢- لنأخذ تجربة إلقاء قطعة نقود معدنية لمَرَّة واحدة فقط، فعندئذ بفرض $p$ هو احتمال ظهور الصورة للأعلى، فعندئذ بوضع $\mathbf{H}_o : p = \frac{1}{2}$، فإنَّ ذلك يعني أنَّه لدينا اختبار للفرضيّة الابتدائية القائلة بتوازن قطعة النقود، وفي حال التوصل إلى رفض هذه الفرضيّة فإنَّه يجب القبول بفرضية أخرى بديلة من أحد الأشكال الآتية وجميعها تُعبِّر عن عدم توازن قطعة النقود:

$$\mathbf{H}_A : p \neq \frac{1}{2} \quad \text{or} \quad \mathbf{H}_A : p > \frac{1}{2} \quad \text{or} \quad \mathbf{H}_A : p < \frac{1}{2}$$

٣- لنفترض أنَّه لدينا مجتمعاً مكوَّناً من درجات الاختبار لطلاب السّنة الأولى المشتركة في جامعة ما، فعندئذ لو وضعنا الفرضيّة الابتدائية:

$$\mathbf{H}_o : F = \mathbf{N}(\mu , \sigma)$$

فإنَّنا نأمل بذلك أن يكون لتوزيع درجات الطلاب توزيع طبيعي بمتوسط $\mu$ وانحراف معياري $\sigma$، وفي حال التوصل إلى رفض هذه الفرضيّة فإنَّه يجب القبول بفرضية أخرى بديلة من الشكل الآتي:

$$\mathbf{H}_A : F \neq \mathbf{N}(\mu , \sigma)$$

نلاحظ من الأمثلة السّابقة أنَّ اختيار الفرضيّة البديلة يتمّ بناءً على طبيعة المسألة المطروحة ومتطلباتها، ولا توجد قاعدة مُحدَّدة تفرض علينا مسبقاً شكليّة الفرضيّة البديلة التي يجب أخذها إلاَّ إذا طلب ذلك صراحة في نص المسألة المُعدّة للدراسة.

**(١١,١,٢) الأخطاء الناتجة عن دراسة فرضيّة إحصائية واحتمالاتها**

نعلم أنَّ قيم الإحصاءات قد تتغيَّر بتغيُّر العيّنات المسحوبة من المجتمع، ومن ثَمَّ قيمة إحصاءة مقدِّرة (كمقدِّر نقطي) لمَعْلَمة مجتمع ما قد تكون قريبة من القيمة المتوقعة للمَعْلَمة في حال صحة الفرضيّة الإحصائية المقدَّمة، وأمَّا إذا كان الفرق ما بين قيمة هذه الإحصاءة والقيمة المتوقعة للمَعْلَمة كبيرة نسبياً، فعندئذ يُقال إنَّ الفرق ذو أهميّة، ونتيجة لذلك تُرفض الفرضيّة الابتدائية $\mathbf{H}_o$، ومن ثَمّ فإنَّه لدى قبول أو رفض فرضيّة إحصائية قد يُرتكب خطأ ما، وهذا الخطأ على نوعين هما:

**الخطأ من النوع الأول Error of the First Type**

إنَّ **الخطأ من النوع الأول** (**ويُعرف باسم Type I Error أيضاً**) هو ذلك الخطأ الناشئ عن رفض الفرضيّة الابتدائية $\mathbf{H}_o$ على الرغم من صحتها.

**الخطأ من النوع الثاني Error of the Second Type**

إنَّ **الخطأ من النوع الثاني** (**ويُعرف باسم Type II Error أيضاً**) هو ذلك الخطأ الناشئ عن قبول الفرضيّة الابتدائية $\mathbf{H}_o$ على الرغم من خطئها.

إنَّ الوقوع في أي من الخطأين السابقين مُمكن باحتمال ما، ولذلك تمَّ تخصيص قيم هذين الاحتمالين بأسماء خاصّة بهما ويقدّمهما لنا التعريفين الآتيين.

**(١ , ٢ , ١ , ١١) تعريف ( مستوى أهمّية الاختبار Significance Level of the Test)**

يُطلق على قيمة احتمال الوقوع في الخطأ من النوع الأول اسم **مستوى أهمّية الاختبار** (**أو المخاطرة - أو المجازفة- من النوع الأول The Risk of Type I**)، ويُرمز لهذه القيمة بـ $\alpha$ .

**(٢ , ٢ , ١ , ١١) تعريف (المخاطرة من النوع الثاني The Risk of Type II)**

يُطلق على قيمة احتمال الوقوع في الخطأ من النوع الثاني اسم **المخاطرة من النوع الثاني**، ويُرمز لهذه القيمة بـ $\beta$ .

**(٣ , ٢ , ١ , ١١) ملاحظات**

١- يُطلق على مستوى أهمّية اختبار فرضيّة إحصائية اسم **مستوى دلالة الاختبار** أيضاً، وأمّا في دراستنا المقبلة فإنّنا سنستخدم عبارة **مستوى الأهمّية** فقط.

٢- يسعى المرء عادةً إلى جعل قيمة $\alpha$ صغيرة، وقد اتفق على أنَّ أخذ $\alpha = 0.05$ مقبولةً إحصائياً؛ ولذلك نرى أنَّ معظم البرامج الإحصائية تنطلق من هذه القيمة لـ $\alpha$ كقيمة افتراضية، ولكن يمكن للمرء التعديل على هذه القيمة حسب متطلبات المسألة المدروسة.

٣- يُطلق على المقدار $1-\beta$ اسم **قوة الاختبار** (**أو حساسية الاختبار**) Power of the Test.

٤- يُطلق على المقدار $1-\alpha$ اسم **مستوى ثقة الاختبار** (**أو احتمال الأمان**)، حيث سنعتمد تسميّة مستوى ثقة الاختبار، وهذه التسميّة قد ورد ذكرها في الفصل السابق عند دراسة فترات الثقة.

٥- يمكن التعبير عن الثابتين $\alpha$ و $\beta$ بوساطة الاحتمالات الشرطية أيضاً، وذلك على النحو الآتي:
بفرض أنَّ $\left[\Omega ,\mathbb{S},\mathcal{P}=\{P_\theta ;\theta \in \Theta\}\right]$ مجتمع موصوف من خلال متغيِّر عشوائي $X$ ، ولنأخذ $\mathfrak{X}_n$ عيّنة من هذا المجتمع، وليكن $T = g\left(\mathfrak{X}_n\right)$ مقدِّراً نقطياً للمَعْلَمة $\theta$ ، وأخيراً لنفترض أنَّ $\mathbf{C}$ فترة من $\mathbb{R}$ ، فعندئذ تكوِّن $\mathbf{C}$ و $\overline{\mathbf{C}}$ تجزئة للمحور الحقيقي $\mathbb{R}$ ، ولنصطلح على الآتي:

أ-نقبل الفرضيّة الابتدائية $\mathbf{H}_o$ إذا كانت قيمة الإحصاءة $T$ تقع في الفترة $\mathbf{C}$ ، ومن ثمَّ رفض $\mathbf{H}_A$ وفي هذه الحالة يُطلق على المنطقة التي تشغلها $\mathbf{C}$ اسم **منطقة القبول** Acceptance Areaللفرضيّة الابتدائية.

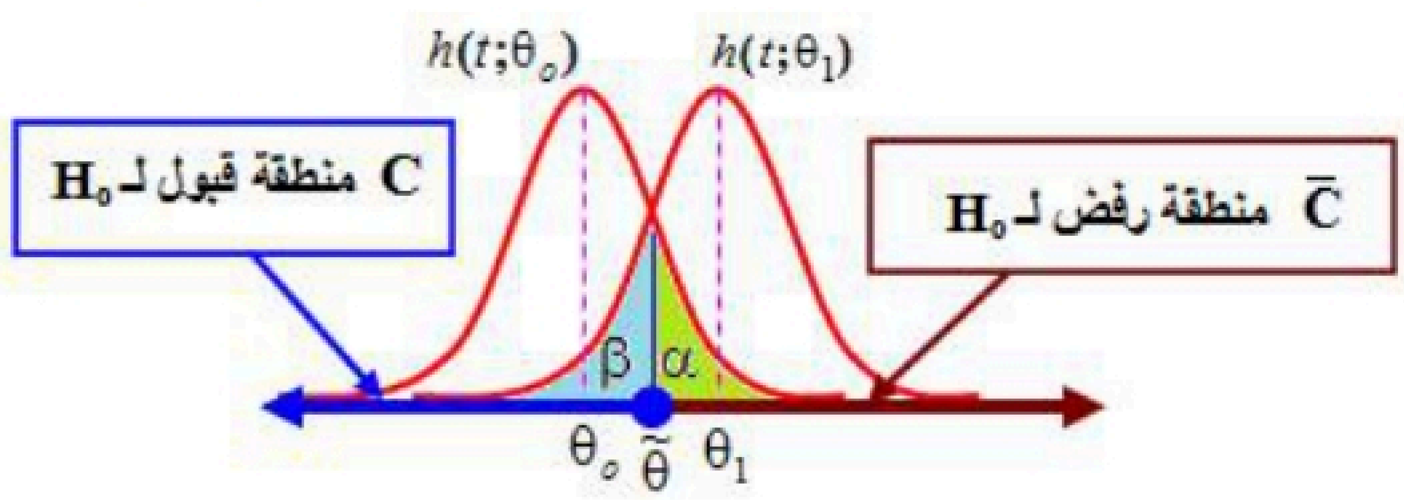
ب- نقبل الفرضيّة البديلة $\mathbf{H}_A$ إذا كانت قيمة الإحصاءة $T$ تقع في $\overline{\mathbf{C}}$ ، ومن ثمَّ رفض $\mathbf{H}_o$ ، وفي هذه الحالة يُطلق على المنطقة التي تشغلها $\overline{\mathbf{C}}$ اسم **منطقة الرفض** Rejection Area للفرضيّة الابتدائية (**أو المنطقة الحرجة للاختبار Critical Region of the Test**).

بناءً على ما تقدَّم يمكننا أن نكتب ما يلي:

$$P\ (T\ \in C\ \mid \mathbf{H}_o\ \ is\ \ false\ ) = \beta \qquad [11,1]$$

$$P\left(T\ \in \bar{C}\ \mid \mathbf{H}_o\ \ is\ \ true\right) = \alpha \qquad [11,2]$$

ولتوضيح العلاقة بين القيمتين $\alpha$ و $\beta$ سنفترض (على سبيل التوضيح والتبسيط) أنَّ للإحصاءة $T$ توزيعًا طبيعيًّا بدالَّة كثافة احتمالية $h\,(t\ ;\ \theta)$، ولنأخذ الاختبار الإحصائي $\mathbf{H}_o : \theta = \theta_o \le \tilde{\theta}$ مقابل الفرضيّة البديلة $\mathbf{H}_A\ :\ \theta = \theta_1 > \tilde{\theta}$، علماً أنَّ قيمة $\tilde{\theta}$ مُثَّبتة وتُدعى **القيمة الحرجة** Critical Value أو **النقطة الحرجة** Critical Point للاختبار، فعندئذ يكون للإحصاءة $T$ تحت الفرضيّة الابتدائية $\mathbf{H}_o$ دالَّة كثافة احتمالية $h\,(t\,;\theta_o)$، وتحت الفرضيّة البديلة $\mathbf{H}_A$ له دالَّة كثافة احتمالية $h\,(t\,;\theta_1)$، وهنا نلاحظ أنَّ الدالَّة $h\,(t\,;\theta_1)$ تنتج عن $h\,(t\,;\theta_o)$ بانسحاب مقداره $\theta_1 - \theta_o$ لأنَّ $t$ ثابتة في هذه الحالة (انظر الشكل الآتي).




الشكل (١ , ١١)

إنَّ السؤال الذي يَطْرح نفسه هنا هو: **ماذا نشترط من اختبار إحصائي ما؟**

في الواقع ما يطمح إليه المرء لدى اختبار إحصائي ما هو معرفة قيمة الإحصاءة $T$ التي من أجلها تكون الفرضيّة الابتدائية $\mathbf{H}_o$ مقبولة، وهذا يعني أنَّه من أجل أيَّة قيمة للعيِّنة $\mathfrak{X}_n$ علينا أنْ نُعطي منطقة قبول ومنطقة رفض للفرضيّة الابتدائية $\mathbf{H}_o$. لذلك ستكون المُهمَّة مركَّزة على كيفية اختيار النقاط على محور المَعْلَمة $\mathbf{o\,\theta}$ (في دراستنا هذه هو المحور الحقيقي نفسه) التي تُحدِّد لنا منطقة القبول ومنطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ التي من أجلها يكون قرار القبول أو الرفض صحيحاً باحتمال معلوم، ولتوضيح ذلك أكثر سنفترض أنَّ $\tilde{\theta}$ هي تلك النقطة التي تفصل منطقة القبول $\mathbf{C}$ عن منطقة الرفض $\bar{\mathbf{C}}$ للفرضيّة $\mathbf{H}_o$، ولذلك دُعيت $\tilde{\theta}$ **النقطة الحرجة** (انظر الشكل السابق (١ , ١١))، علاوةً على ذلك سنفترض أنَّه من مُعطيات العيِّنة $\mathfrak{X}_n$ حصلنا على قيمة $t'$ للإحصاءة $T$ بحيث إنَّ $\tilde{\theta} < t'$، أي إنَّ $t' \ni \bar{\mathbf{C}}$، فعندئذ تكون الفرضيّة الابتدائية $\mathbf{H}_o$ مرفوضة، ولكنَّ هذا القرار قد يكون غير صحيح لأنَّ القيمة $t'$ يمكن أنْ نحصل عليها عندما تكون $\theta \le \tilde{\theta}$ أيضاً، ومن ثَمَّ يكون لدينا ما يلي:

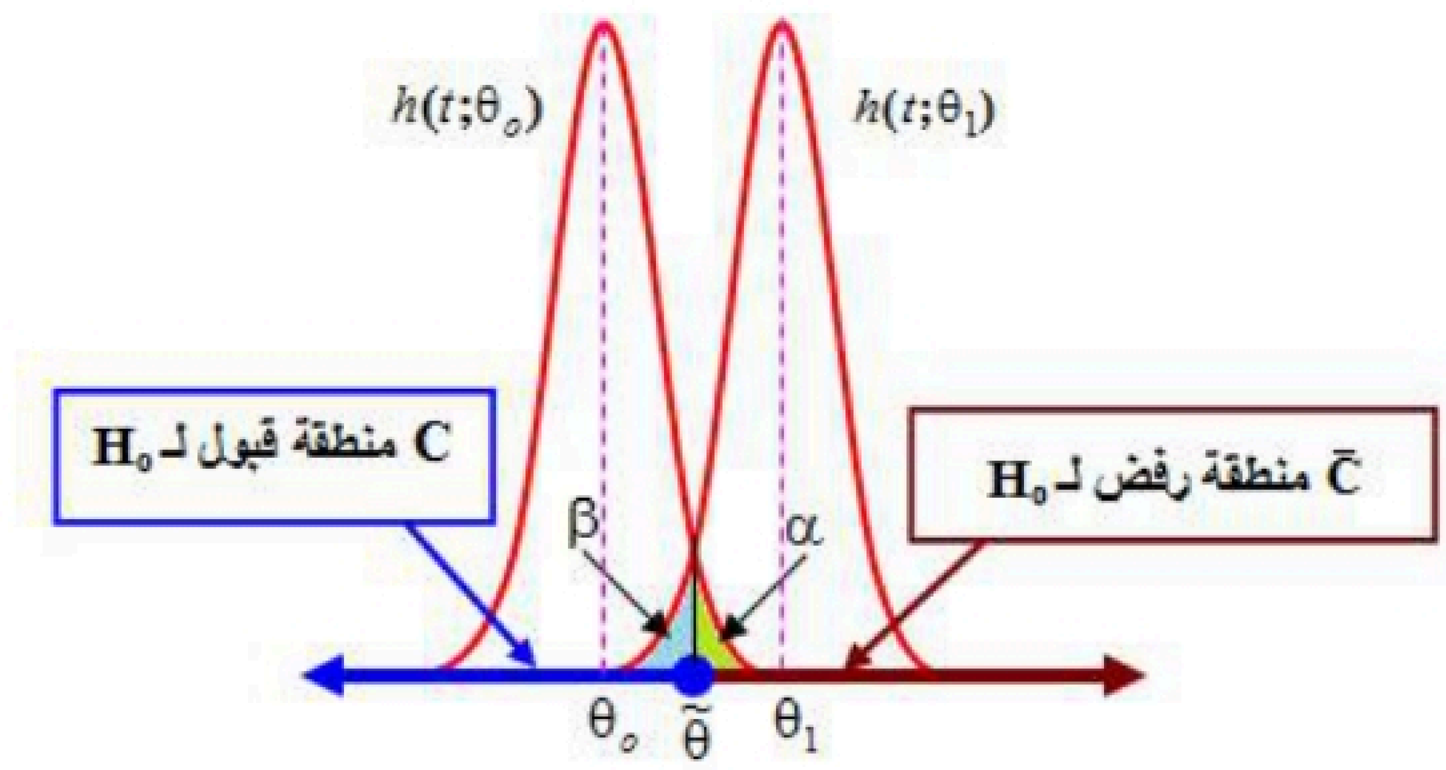
$$\alpha = P\,(T\ \in \bar{\mathbf{C}} \mid \mathbf{H}_o\ \ is\ \ true\,) = \int_{\tilde{\theta}}^{\infty} h\,(t\ ;\ \theta_o\,)\ dt \ge \int_{\tilde{\theta}}^{\infty} h\,(t'\,;\ \theta_o\,)\ \ dt' \qquad [11,3]$$

$$\beta = P\,(T\ \in \mathbf{C}\,|\mathbf{H}_o\ \ is\,false\,) = \int_{-\infty}^{\tilde{\theta}} h\,(t\,;\theta_1)\,dt = 1 - \int_{\tilde{\theta}}^{\infty} h\ \,(t\,;\theta_1)\,dt \qquad [11,4]$$

وهكذا يُلاحظ أنَّه من أجل حجم عيِّنة مثَّبت (ومن ثَمَّ تفرطح $h\left(t\ ;\ \theta\right)$ ثابت)، فإنَّ أي زيادة في قيمة $\alpha$ سيرافقها نقصان في قيمة $\beta$ والعكس صحيح أيضاً (انظر الشكل السابق (١ , ١١)). من جانب آخر نلاحظ أنَّ $\alpha$ هي أكبر قيمة احتمال تكون من أجلها $\mathbf{H}_o$ مرفوضة، ولهذا السبب فكلما كان مستوى الأهمّيّة $\alpha$ أصغر أدى ذلك إلى اختبار أفضل للفرضيّة الابتدائية $\mathbf{H}_o$، ولكن كما ذكرنا قبل قليل أنَّه في معظم الحالات يكون الاختبار مقبولاً إحصائياً عند مستوى من الأهمّيّة قدره $\alpha \le 0.05$، وهذا غالباً ما يستخدم في التطبيقات العمليّة.

**الآن لننظر ما الذي سيحصل عندما نزيد حجم العيّنة؟**

إذا قمنا بزيادة حجم العيّنة $\mathcal{X}_n$ على ما كانت عليه سابقاً، فعندئذ تباين الإحصاءة $T$ سوف يتناقص، وذلك لأنَّ تباين هذه الإحصاءة متناسب عكساً مع حجم العيّنة (بحسب توزيعات المعاينة التي درسناها سابقاً)، ومن ثمَّ فإنَّ دالَّة الكثافة لهذه الإحصاءة سوف تصبح أكثر تدبباً وأطرافها أقل انبساطاً (انظر الشكل الآتي)، وبالتالي ينتج عنه تناقص في قيمتي $\sigma$ و $\beta$ معاً من أجل قيمة مُثَبّتة لـ $\tilde{\theta}$.




الشكل (٢ , ١١) قارن هذا العرض مع الشكل (١ , ١١)

الجدول الآتي يقدِّم لنا القرارات المتخذة لدى اختبار إحصائي ما:

الجدول (١ , ١١)

| | | اتخاذ القرار | نتيجة القرار | احتمال نتيجة القرار يساوي |
|---|---|---|---|---|
| الحال الحقيقي للفرضيّة | $H_o$ صحيحة | $H_o$ مقبولة<br>$H_A$ مرفوضة | قرار صحيح | $1-\alpha$<br>مستوى ثقة الاختبار |
| | | $H_o$ مرفوضة<br>$H_A$ مقبولة | قرار خاطئ<br>(خطأ من النوع الأول) | $\alpha$ ستوى الأهميّة<br>(أو المخاطرة من النوع الأول) |
| | $H_o$ خاطئة | $H_o$ مقبولة<br>$H_A$ مرفوضة | قرار خاطئ<br>(خطأ من النوع الثاني) | $\beta$<br>(أو المخاطرة من النوع الثاني) |
| | | $H_o$ مرفوضة<br>$H_A$ مقبولة | قرار صحيح | $1-\beta$<br>قوة الاختبار |

**(٤ , ٢ , ١ , ١١) ملاحظات**

١- بما أنَّنا نهتم عادةً بعدم الرفض الخاطئ لفرضيّة إحصائية، فإنَّه علينا أن نأخذ الفرضيّة الابتدائية بحيث يكون احتمال رفضها على الرغم من صحتها صغيراً، وغالباً يؤخذ احتمال رفضها بحيث يكون أصغر أو يساوي 0.05.

٢- من أجل اختبار مُحدَّد وعيِّنة مثبتة الحجم لا يمكن تصغير $\alpha$ و $\beta$ معاً، ولكن يُمكننا أن نُصغِّر قيمة $\alpha$ بالقدر الذي نريد عندما لا نكون مهتمين بقيمة $\beta$.

٣- بتكبير حجم العيّنة $n$ نستطيع تصغير كل من $\alpha$ و $\beta$ معاً، ومن ثَّم تحسين قوة الاختبار $1-\beta$.

لقد ذكرنا سابقاً أنَّه يمكن تصنيف الفرضيّات الإحصائية إلى نوعين أساسيين هما **الاختبارات المَعْلميّة** (أي عندما يكون الاختبار

متعلّق بمَعَالم المجتمع) و **الاختبارات غير المَعْلميّة** (أي عندما يكون الاختبار متعلّق بأي شيء يخصّ المجتمع عدا مَعَالمه). أكثر من ذلك سنلاحظ فيما يلي أنَّ اختبارات الفرضيّات الإحصائية المَعْلميّة يُمكن أن تصنَّف في نوعين رئيسيين أيضاً يُقدِّمها لنا التعريفين الآتيين.

**(١١, ١, ٢, ٥) تعريف (الفرضيّة المركّبة Composite hypothesis)**

ليكن $[\Omega, \mathbb{S}, \mathcal{P}=\{P_\theta;\theta\in\Theta\}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ ، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، فإذا كان اختبار الفرضيّة الابتدائية من النموذج $\mathbf{H}_o : \theta \in \Theta_o$ مقابل فرضيّة بديلة من النموذج $\mathbf{H}_A : \theta \notin \Theta_o$ علماً أنَّ $\mathbb{R}^k \supseteq \Theta \supseteq \Theta_o$ و $k \in \mathbb{N}$ ، فعندئذ يُقال إنّه لدينا اختبار **فرضيّة مركّبة**.

**(١١, ١, ٢, ٦) تعريف (الفرضيّة البسيطة Simple hypothesis)**

ليكن $[\Omega, \mathbb{S}, \mathcal{P}=\{P_\theta;\theta\in\Theta\}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ ، ولنأخذ $\mathfrak{X}_n$ عيِّنة من هذا المجتمع، فإذا كان اختبار الفرضيّة الابتدائية من النموذج $\mathbf{H}_o : \theta = \theta_o$ مقابل فرضيّة بديلة من أحد النماذج الآتية:

$$\mathbf{H}_A : \theta \neq \theta_o \quad \text{or} \quad \mathbf{H}_A : \theta > \theta_o \quad \text{or} \quad \mathbf{H}_A : \theta < \theta_o$$

فعندئذ يُقال إنَّه لدينا اختبار **فرضيّة بسيطة**.

**(١١, ١, ٢, ٧) ملاحظة**

إذا كانت $\theta \in \Theta \subseteq \mathbb{R}^k$ ، وبفرض أنَّ الاختبار مُتعلّق ببعض مركّبات المتّجه $\theta = (\theta_1, \theta_2, ..., \theta_k)$ فإنَّ بقية المركّبات التي لم يشملها الاختبار تُدعى **مَعَالم مشوّشة** Perturbation Parameters.

**(١١, ١, ٢, ٨) أمثلة**

١- ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً بواسونياً بمَعْلَمة $\lambda > 0$ ، فإذا أخذنا الفرضيّة الابتدائية $\mathbf{H}_o : \lambda = 2$ مقابل إحدى الفرضيّات البديلة $\mathbf{H}_A : \lambda \neq 2$ أو $\mathbf{H}_A : \lambda > 2$ أو $\mathbf{H}_A : \lambda < 2$ ، ونكون عندئذ أمام فرضيّة بسيطة (ليس فيها مَعَالم مشوشة).

٢- ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً حدَّانياً بمَعْلمتين $n$ و $p$ ، فإذا أخذنا الفرضيّة الابتدائية $\mathbf{H}_o : p = 0.5$ مقابل إحدى الفرضيّات البديلة $\mathbf{H}_A : p \neq 0.5$ أو $\mathbf{H}_A : p > 0.5$ أو $\mathbf{H}_A : p < 0.5$ ، ونكون عندئذ أمام اختبار لفرضيّة بسيطة فيها المَعْلَمة $n$ هي مَعْلَمة مشوّشة.

٣- ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً طبيعياً بمتوسط $\mu$ وانحراف معياري $\sigma$ ، فإذا أخذنا الفرضيّة الابتدائية $\mathbf{H}_o : \mu < 5$ التي تكافئ الفرضيّة الابتدائية $\mathbf{H}_o : (\mu, \sigma) \in (-\infty, 5) \times (0, \infty)$ مقابل الفرضيّة البديلة $\mathbf{H}_o : \mu \geq 5$ التي تكافئ الفرضيّة البديلة $\mathbf{H}_A : (\mu, \sigma) \in [5, \infty) \times (0, \infty)$ ، فإنَّنا سنكون أمام اختبار لفرضيّة مركّبة.

**(١١, ١, ٢, ٩) ملاحظات**

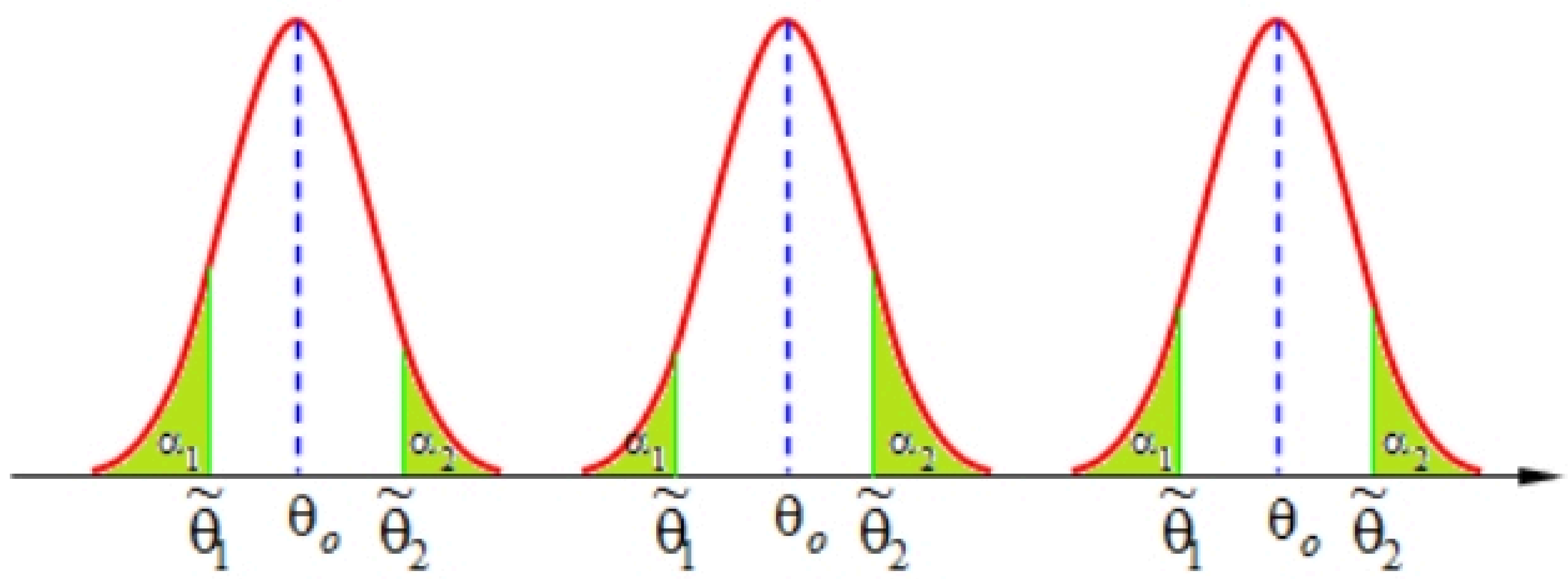
١- إذا كان لدينا اختبار لفرضيّة ابتدائية من النموذج $\mathbf{H}_o : \theta = \theta_o$ مقابل فرضيّة بديلة من النموذج $\mathbf{H}_A : \theta \neq \theta_o$ ، فعندئذ يُقال إنَّه لدينا **اختبار ثنائي الطرف** (أو ثنائي الذيل) Two-Tailed Test ، وأمّا إذا كانت الفرضيّة البديلة من النموذج $\mathbf{H}_A : \theta > \theta_o$ أو $\mathbf{H}_A : \theta < \theta_o$ ، فعندئذ يُقال إنّه لدينا **اختبار أحادي الطرف** (أو أحادي الذيل) One -Tailed Test.

٢- في هذا الفصل من الكتاب لن نقوم بدراسة الفرضيّات المُركّبة.

### (١١, ١, ٣) تعيين منطقة الرفض المثلى للفرضيّة الابتدائية

**Determination the Optimum Rejection Area of Primary Hypothesis**

لقد لاحظنا خلال توضيحنا للمقدارين $\alpha$ و $\beta$ أنّنا اخترنا نموذجاً لاختبار وحيد الطرف، وقد كانت المساحة التي تُحدِّد منطقة الرفض (وقدرها $\alpha$) تقع في جهة واحدة، وبالتالي من أجل حجم عيّنة مُثبَّت ستكون منطقة الرفض مُحدَّدةً تماماً. لكن لو كان لدينا اختبار ثنائي الطرف، فعندئذ ستكون هذه المساحة الممثِّلة لمنطقة الرفض مجزأة على طرفي دالّة الكثافة الاحتمالية بحيث يكون عند الطرف الأيسر مساحة قدرها $\alpha_1$ وعند الطرف الأيمن مساحة قدرها $\alpha_2$ وبحيث يكون $\alpha_1 + \alpha_2 = \alpha$، أي سيكون لدينا أحد العروض الموجودة في الشكل الآتي (٣, ١١) (وجميعها تعالج الاختبار نفسه).




الشكل (٣, ١١)

فنلاحظ إمكانية وجود أكثر من منطقة رفض ناتجة عن القيمة $\alpha$ نفسها، ومن ثمّ منطقة القبول للفرضيّة الابتدائية $\mathbf{H}_o$ قد تطول حيناً وتقصر حيناً آخر، ولذلك يطرح السؤال الآتي:

**أي من هذه المناطق سنختار كمنطقة قبول للفرضيّة الابتدائية؟**

تُعطينا النظرية الآتية (التي سنقبلها دون برهان) أفضل منطقة رفض للفرضيّة الابتدائية $\mathbf{H}_o : \theta = \theta_o$ مقابل الفرضيّة البديلة $\mathbf{H}_A : \theta \neq \theta_o$.

### (١١, ١, ٣, ١) نظرية (نيمان – بيرسون Neyman Pearson Theorem)

ليكن $\left[\Omega, \mathcal{S}, \mathcal{P}=\{P_\theta ; \theta \in \Theta\}\right]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$، ولنأخذ $\mathcal{X}_n$ عيّنة من هذا المجتمع، ولنفترض أنّنا نريد اختبار فرضيّة ابتدائية $\mathbf{H}_o : \theta = \theta_o$ مقابل فرضيّة بديلة $\mathbf{H}_A : \theta \neq \theta_o$ عند مستوى من الأهمّيّة قدره $\alpha$. تحت هذه الفرضيّات تُعطى **منطقة الرفض المُثلى** Optimal Rejection Area (وسنرمز لها بـ $\mathbf{C}^*$) للفرضيّة الابتدائية $\mathbf{H}_o$ من خلال المجموعة الآتية:

$$\mathbf{C}^* = \left\{ \vec{x}_n \ ; \ \frac{L(\vec{x}_n ; \theta_1)}{L(\vec{x}_n ; \theta_o)} \geq k \right\} \qquad [11,5]$$

علماً أنَّ $L(\vec{x}_n ; \theta)$ هي دالّة الأرجحية للعيِّنة $\mathcal{X}_n$ ومُعطاة بالعلاقة [10-2]، وأمّا $\vec{x}_n$ فهي قيمة الملحوظة للعيِّنة $\mathcal{X}_n$، وأخيراً $0 < k$ ثابت يُعيَّن من خلال العلاقة الآتية:

$$P\left(\mathcal{X}_n \in \mathbf{C}^* \mid \mathbf{H}_o \ is \ true\right) = \alpha \qquad [11,6]$$

### (١١, ١, ٣, ٢) ملاحظة

إنَّ الاختبار الذي يعتمد على منطقة رفض مُثلى وفقاً لنظرية نيمان– بيرسون يُدعى **أفضل اختبار** Best Test أو **الاختبار**

**الأمثل** Optimal Test ، ويُبرهن في هذه الحالة أنَّ أقصر طول للفترة التي تُعيِّن منطقة القبول للفرضيّة الابتدائية $\mathbf{H}_o$ والموافقة لأصغر خطأ من النوع الثاني $\beta$ هي الفترة التي يقع على يمينها $\frac{\alpha}{2}$ وعلى يسارها $\frac{\alpha}{2}$ من مساحة حجم منطقة الرفض.

**مثال (١١,١,٣,٣)**

ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً طبيعياً بمتوسط $\mu$ وانحراف معياري $\sigma = \sigma_o$ معلوم، ولنأخذ $\mathcal{X}_n$ عيّنة من هذا المجتمع، ولنقم بتحديد منطقة الرفض المُثلى للفرضيّة الابتدائية $\mathbf{H}_o\ : \mu = \mu_o$ مقابل الفرضيّة البديلة $\mathbf{H}_A\ : \mu = \mu_1 \neq \mu_o$ ، وذلك عند مستوى من الأهميّة قدره $\alpha$ .

**الحل**: نعلم أنَّ دالَّة الأرجحيَّة لعيِّنة عشوائية $\mathcal{X}_n$ مسحوبة من مجتمع طبيعي لها العرض الآتي:

$$L\left(\vec{x}_n ; (\mu, \sigma_o)\right) = \left(\frac{1}{2\pi\sigma_o^2}\right)^{\frac{n}{2}} \cdot \mathbf{exp}\left[-\frac{1}{2\sigma_o^2}\sum_{k=1}^{n}\left(x_k - \mu\right)^2\right]$$

ومن أجل $\mu_1 \neq \mu_o$ يكون لدينا ما يلي:

$$\frac{L\left(\vec{x}_n ; (\mu_1, \sigma_o)\right)}{L\left(\vec{x}_n ; (\mu_o, \sigma_o)\right)} = \mathbf{exp}\left[\frac{\mu_1 - \mu_o}{\sigma_o^2}\sum_{k=1}^{n} x_k + \frac{n\left(\mu_o^2 - \mu_1^2\right)}{2\,\sigma_o^2}\right]$$

ومن ثَمَّ تُحدَّد منطقة الرفض المُثلى $\mathbf{C}^*$ لهذا الاختبار من خلال العلاقة الآتية:

$$\mathbf{C}^* = \left\{\vec{x}_n\ ;\ \mathbf{exp}\left[\frac{\mu_1 - \mu_o}{\sigma_o^2}\sum_{k=1}^{n} x_k + \frac{n\left(\mu_o^2 - \mu_1^2\right)}{2\,\sigma_o^2}\right] \geq k\right\}$$

علماً أنَّ $k$ ثابت يجب تعيينه، ومن أجل ذلك سنقدِّم المناقشة الآتية:

بما أنَّ $\mu_1 \neq \mu_o$ ، فعندئذ يُمكننا أن نفترض أنَّ $\mu_1 > \mu_o$ ، ومنه سيكون وفقاً لمعطيات المجموعة $\mathbf{C}^*$ ما يلي مُحقَّقاً:

$$\sum_{k=1}^{n} x_k \geq \frac{\sigma_o^2 \ln k}{\mu_1 - \mu_o} + \frac{n\left(\mu_1 + \mu_o\right)}{2} \triangleq a \qquad \mathbf{(1)}$$

فنلاحظ أنَّ $a \in \mathbb{R}$ وذلك لأنَّ $k \in (0, \infty)$ ، ومن ثَمَّ يصبح لـ $\mathbf{C}^*$ العرض $\mathbf{C}^* = \left\{\vec{x}_n\ ;\ \sum_{k=1}^{n} x_k \geq a\right\}$ ، ولنقم بتعين القيمة $a$ بحيث تكون العلاقة [11,6] مُحقَّقة، حيث لدينا:

$$P\left(\mathcal{X}_n \in \mathbf{C}^* \mid \mathbf{H}_o \textit{ is true}\right) = P\left(\mathcal{X}_n \in \mathbf{C}^* \mid \mu = \mu_o\right) = \alpha$$

وباستخدام العلاقة **(1)** يُمكننا أن نكتب الآتي:

$$\alpha = P\left(\mathcal{X}_n \in \mathbf{C}^* \mid \mu = \mu_o\right) = P\left(\sum_{k=1}^{n} x_k \geq a \mid \mu = \mu_o\right) = P\left(n \cdot \bar{X} \geq a \mid \mu = \mu_o\right)$$

$$= P\left(\bar{X} \geq \frac{a}{n} \mid \mu = \mu_o\right) = P\left(\frac{\bar{X} - \mu_o}{\frac{\sigma_o}{\sqrt{n}}} \geq \frac{\frac{a}{n} - \mu_o}{\frac{\sigma_o}{\sqrt{n}}}\right)$$

والآن بوضع $Z := \frac{\bar{X} - \mu_o}{\sigma_o}\sqrt{n}$ وكذلك $c := \frac{\frac{a}{n} - \mu_o}{\sigma_o}\sqrt{n}$ ينتج لدينا أنَّ $P(Z \geq c) = \alpha$ ، ومن ثَمّ يكون لدينا

$P(Z < c) = 1-\alpha$ ، وبما أنَّ $Z \sim \mathbf{N}(0,1)$ ، فإنَّه سيكون لـ $c$ القيمة $c = z_{1-\alpha}$ ، ومن ثمَّ ينتج أنَّ قيمة $a$ تساوي:

$$a = n \cdot \mu_o + z_{1-\alpha}\, \sigma_o \sqrt{n}$$

وهكذا يصبح لـ $k$ القيمة الآتية:

$$k = \exp\left[\frac{\mu_1 - \mu_2}{\sigma_o^2} \cdot a - \frac{n \cdot \left(\mu_1^2 - \mu_o^2\right)}{2\sigma_o^2}\right]$$

ومنه يكون للمنطقة $\mathbf{C}^*$ العرض الآتي:

$$\mathbf{C}^* = \left\{ \vec{x}_n \ ;\ \sum_{k=1}^{n} x_k \geq \alpha = n \cdot \mu_o + z_{1-\alpha} \sigma_o \sqrt{n} \right\}$$

الآن، وبشكل مماثل لما سبق نجد أنَّه إذا كان $\mu_o > \mu_1$ فإنَّ منطقة الرفض المُثلى الموافقة لمستوى أهمّيّة قدره $\alpha$ سيكون لها العرض الآتي:

$$\mathbf{C}^* = \left\{ \vec{x}_n \ ;\ \sum_{k=1}^{n} x_k \leq \alpha = n \cdot \mu_o + z_{1-\alpha} \sigma_o \sqrt{n} \right\}$$

أخيراً إذا كان $\mu_1 \neq \mu_o$ دون تحديد أي من $\mu_o$ و $\mu_1$ هو أصغر من الآخر، فإنَّ منطقة الرفض المُثلى الموافقة لمستوى أهمّيّة قدره $\alpha$ سيكون لها العرض الآتي:

$$\mathbf{C}^* = \left\{ \vec{x}_n \ ;\ \sum_{k=1}^{n} x_k \leq \alpha = n \cdot \mu_o + z_{1-\frac{\alpha}{2}} \sigma_o \sqrt{n} \right\} \cup \left\{ \vec{x}_n \ ;\ \sum_{k=1}^{n} x_k \geq \alpha = n \cdot \mu_o + z_{1-\frac{\alpha}{2}} \sigma_o \sqrt{n} \right\}$$

وذلك لأنَّ منطقة الرفض في هذه الحالة تكون منطقة معيَّنة بالمتباينتين $z \geq -z_{1-\frac{\alpha}{2}}$ و $z \leq z_{1-\frac{\alpha}{2}}$ ، وهاتين العلاقتين تُعطيان أقصر طول للفترة التي تُعيِّنمنطقة القبول للفرضيّة الابتدائية $\mathbf{H}_o : \mu = \mu_o$ .

**ملاحظة (١١ , ١ , ٣ , ٤)**

عندما يرد ذكر منطقة الرفض لفرضيّة ابتدائية $\mathbf{H}_o$ إنَّما نعني بذلك منطقة الرفض المُثلى دون الخوض في إثبات ذلك.

## (١١,٢) الاختبارات الإحصائية من أجل معلمتي مجتمع طبيعي

**Statistical Tests for Parameters of Normal Population**

فيما يلي سنركِّز دراستنا على الاختبارات المَعْلمّية لمجتمعات طبيعية (أو طبيعية على وجه التقريب حيث سنفترضها طبيعية أيضاً)، ومن أجل ذلك ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع للتوزيع $\mathbf{N}(\mu, \sigma)$ ، ولنأخذ $\mathcal{X}_n$ عيّنة من هذا المجتمع بمتوسط $\bar{x}$ وانحراف معياري $s$ ، وعلى سبيل التذكير فإنَّ $\bar{x}$ و $s$ هما قيمتي الإحصاءتين $\bar{X}$ و $S$ على الترتيب.

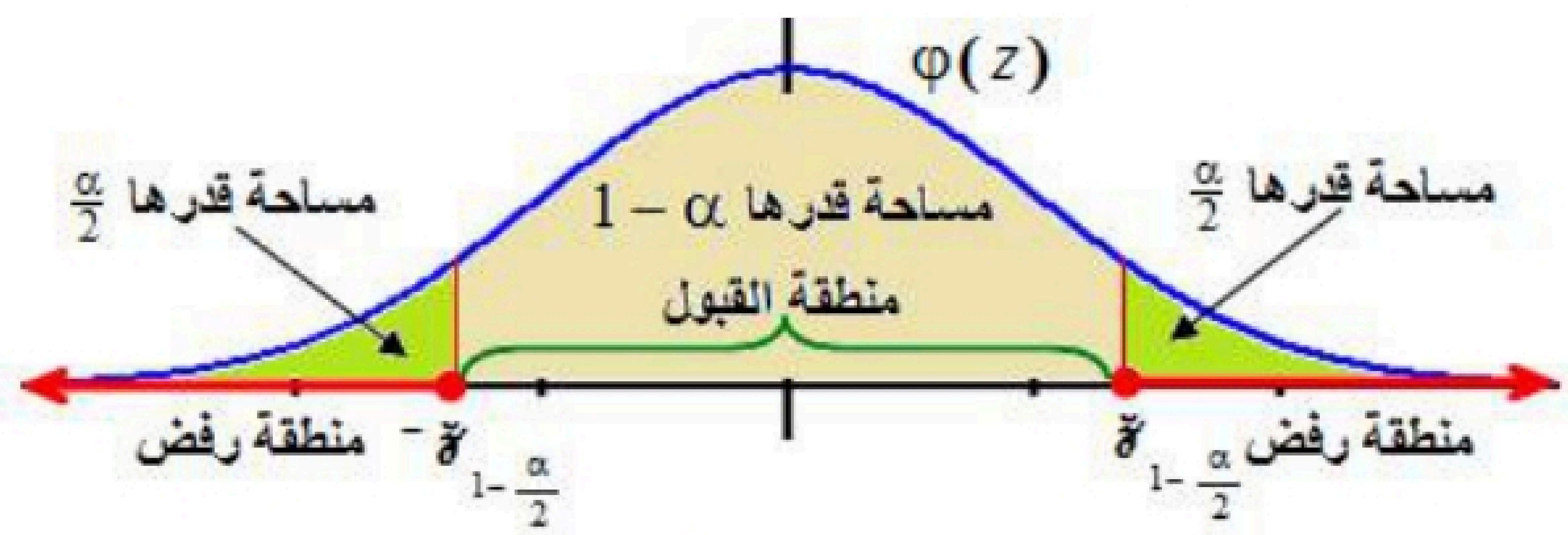
### (١ , ٢ , ١١) اختبارات من أجل متوسط مجتمع طبيعي

تُقدِّم لنا المبرهنات الآتية (سنقبلها دون برهان) اختبارات الفرضيّات المتعلِّقة بـ $\mu$ متوسط المجتمع الطبيعي.

**(١١,٢,١,١) مبرهنة**

تحت الفرضيّات المقدَّمة آنفًا وبفرض أنَّ الانحراف المعياري للمجتمع $\sigma = \sigma_o$ معلوماً، وأنَّه لدينا الفرضيّة الابتدائية $\mathbf{H}_o\ : \mu = \mu_o$، فعندئذ:

١- إذا كانت الفرضيّة البديلة هي $\mathbf{H}_A\ : \mu \neq \mu_o$، فإنَّ منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ وعند مستوى من الأهمّية قدره $\alpha$ تكون معيَّنةً بالمنطقة المحدَّدة بالمتباينة $|z| \geq z_{1-\frac{\alpha}{2}}$ (انظر الشكل الآتي).

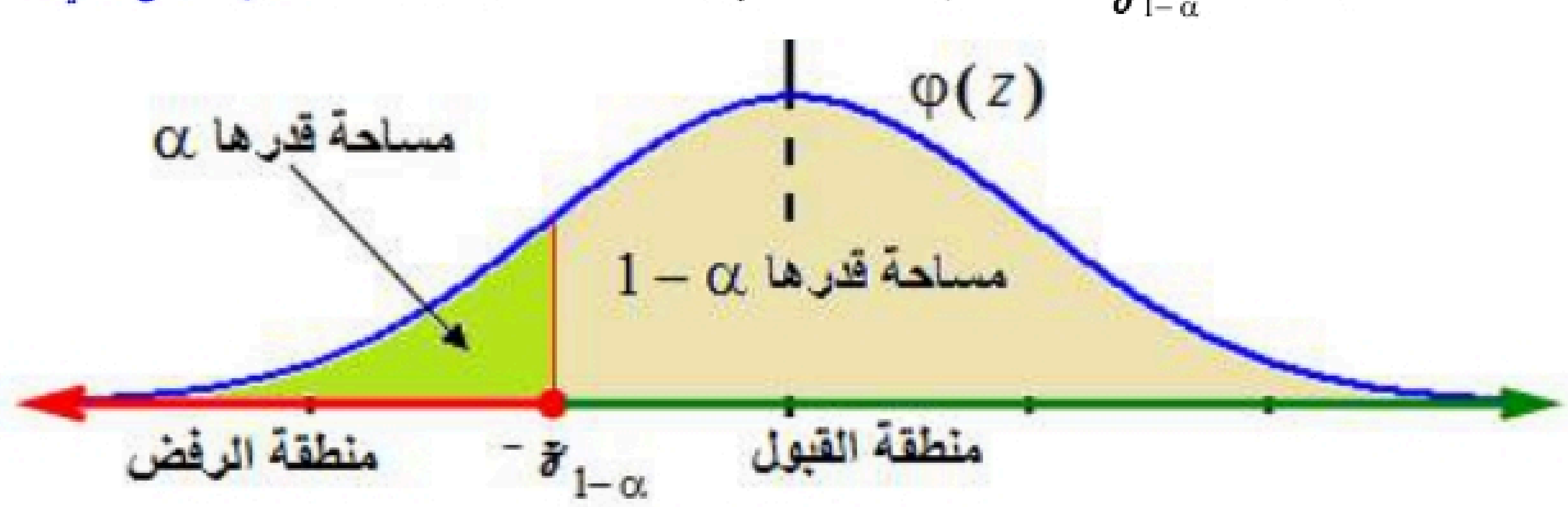



الشكل (٤, ١١. أ)

علماً أنَّ $z$ تُحسب بالعلاقة الآتية:

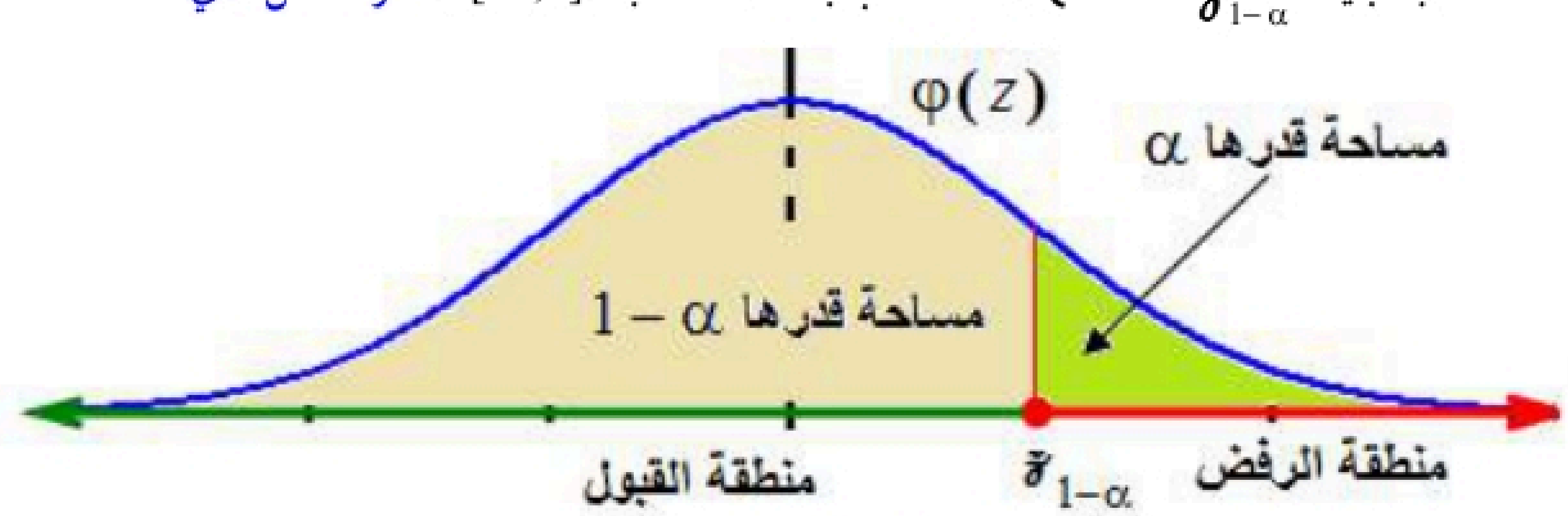
$$z = \frac{\overline{x} - \mu_o}{\sigma_o}\sqrt{n} \qquad [11,7]$$

٢- إذا كانت الفرضيّة البديلة هي $\mathbf{H}_A\ : \mu < \mu_o$، فإنَّ منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ وعند مستوى من الأهمّية قدره $\alpha$ تكون معيَّنةً بالمنطقة المحدَّدة بالمتباينة $z \leq -z_{1-\alpha}$ علماً أنَّ $z$ تُحسب بالعلاقة السابقة [11,7] (انظر الشكل الآتي).




الشكل (٤, ١١. ب)

٣- إذا كانت الفرضيّة البديلة هي $\mathbf{H}_A\ : \mu > \mu_o$، فإنَّ منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ وعنـد مستوى من الأهمّية قدره $\alpha$ تكون معيَّنةً بالمنطقة المحدَّدة بالمتباينة $z \geq z_{1-\alpha}$ علماً أنَّ $z$ تُحسب بالعلاقة السابقة [11,7] (انظر الشكل الآتي).



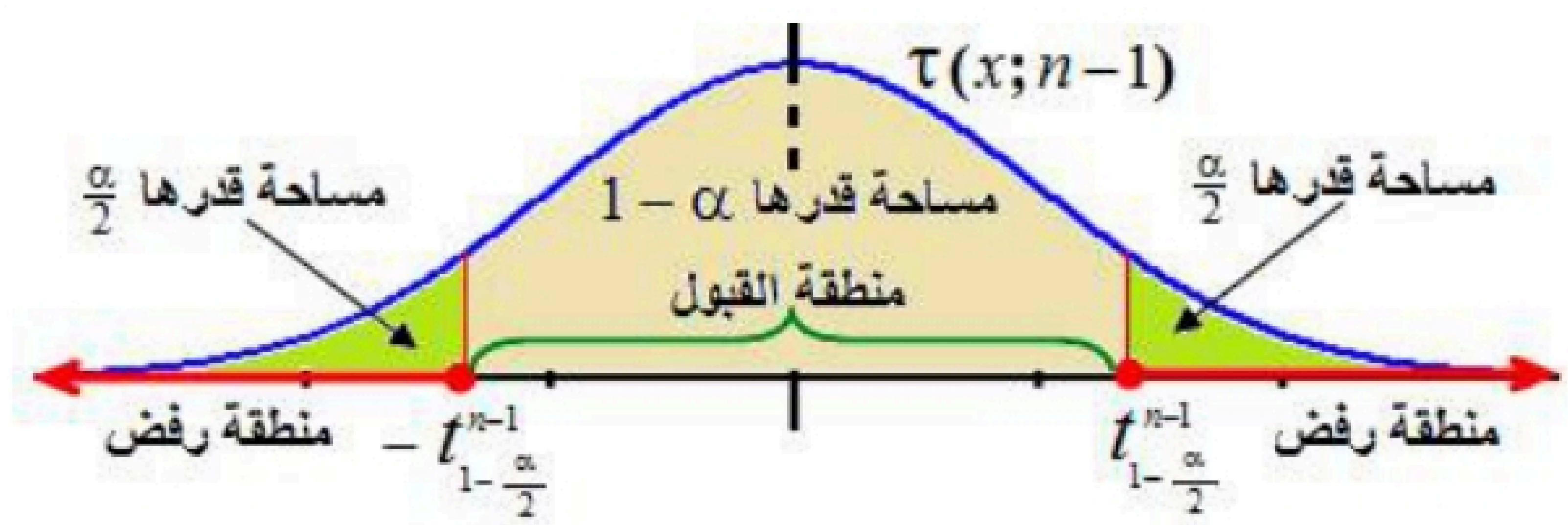

الشكل (٤, ١١. ج)

**(١١,٢,١,٢) مبرهنة**

تحت الفرضيّات المقدَّمة آنفًا وبفرض أنَّ الانحراف المعياري للمجتمع مجهولاً (أي إنَّ $\sigma = ?$)، وأنّه لدينا الفرضيّة الابتدائية $\mathbf{H}_o : \mu = \mu_o$، فعندئذ:

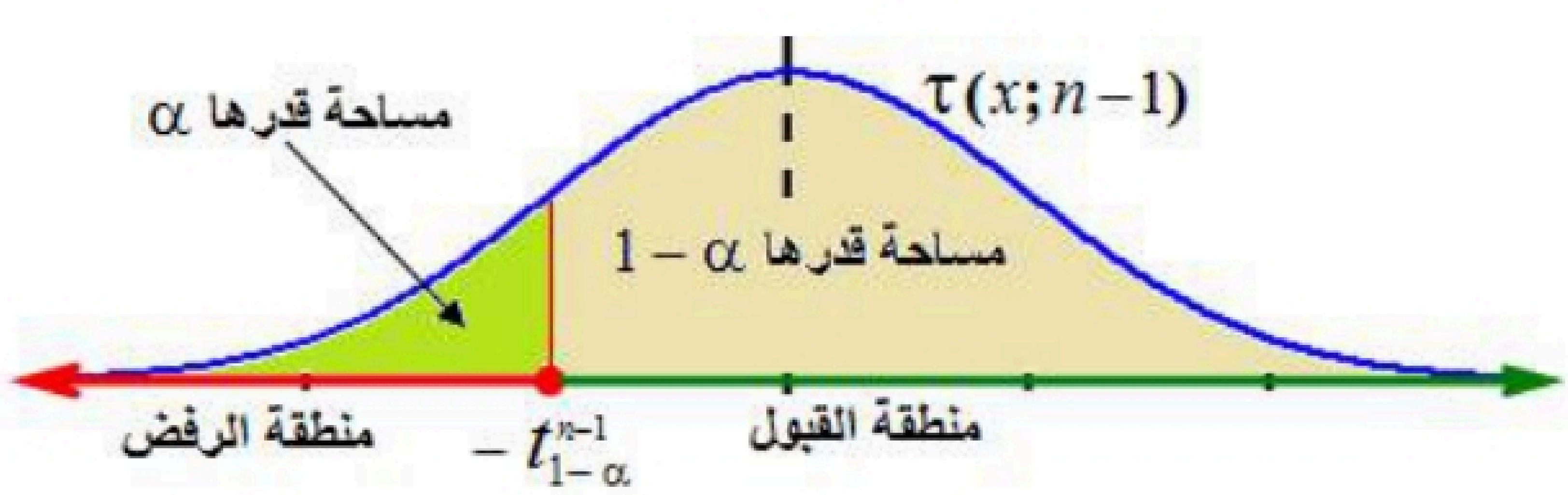
١- إذا كانت الفرضيّة البديلة هي $\mathbf{H}_A : \mu \neq \mu_o$، فإنَّ منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ وعند مستوى من الأهميّة $\alpha$ تُعيَّن بالمنطقة المحدَّدة بالمتباينة $|t| \geq t^{[n-1]}_{1-\frac{\alpha}{2}}$ (انظر الشكل الآتي) علماً أنَّ $t$ تُحسب بالعلاقة الآتية:

$$t = \frac{\bar{x} - \mu_o}{s}\sqrt{n} \qquad [11,8]$$



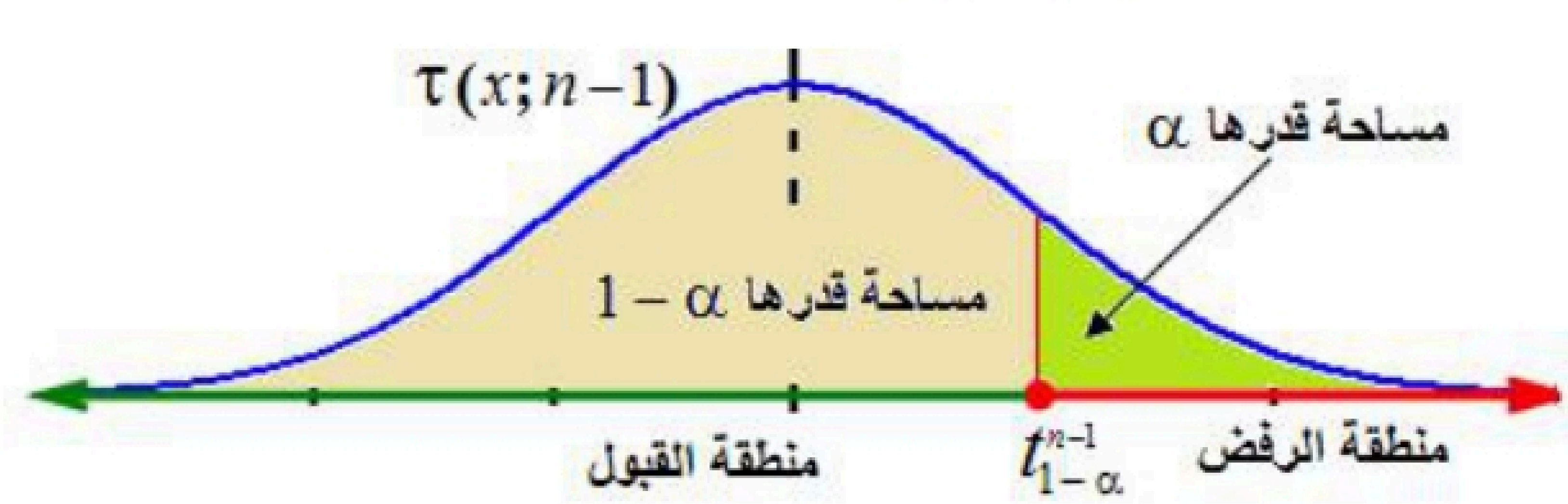

الشكل (١١,٥.أ)

٢- إذا كانت الفرضيّة البديلة هي $\mathbf{H}_A : \mu < \mu_o$، فإنَّ منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ وعند مستوى من الأهميّة $\alpha$ تُعيَّن بالمنطقة المحدَّدة بالمتباينة $t \leq -t^{[n-1]}_{1-\alpha}$ (انظر الشكل الآتي) علماً أنَّ قيمة $t$ تُحسب بالعلاقة السابقة [11,8].




الشكل (١١,٥.ب)

٣- إذا كانت الفرضيّة البديلة هي $\mathbf{H}_o : \mu > \mu_o$، فإنَّ منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ وعند مستوى من الأهميّة $\alpha$ تُعيَّن بالمنطقة المحدَّدة بالمتباينة $t \geq t^{[n-1]}_{1-\alpha}$ (انظر الشكل الآتي) علماً أنَّ $t$ تُحسب بالعلاقة السابقة [11,8].




الشكل (١١,٥.ج)

**(٣,١,٢,١١) مثال**

زعم مصنع لإنّتاج بطاريات 3.7 فولت للهاتف المحمول أنَّ متوسط عمل البطاريات لديه بعد شحنها بنسبة 100% وعند سحب تيار ثابت قدره 20 ميلي أمبير يساوي 72 ساعة بانحراف معياري قدره $\sigma_o$ ساعة، قمنا بأخذ عيّنة من إنتاج هذا المصنع بحجم $n$ بطارية فوجدنا أنَّ متوسط عمل هذه البطاريات وتحت الفرضيّات المذكورة آنفًا يساوي 70 ساعة بانحراف معياري قدره 4 ساعات، فإذا افترضنا أنَّ أعمار شحن البطاريات يخضع للتوزيع الطبيعي، فعندئذ من أجل مستوى أهميّة قدره $\alpha = 0.05$ لنقم بما يلي:

١- اختبار الفرضيّة القائلة إنَّ متوسط عمل البطارية يساوي 72 ساعة تحت الفرضيّات المذكورة آنفًا علماً أنَّ $\sigma_o = 5$ ساعات و $n = 36$.

٢-إعادة حل الطلب السابق إذا كان مستوى الأهميّة $\alpha = 0.01$.

٣-اختبار الفرضيّة القائلة إنَّ متوسط عمل البطارية يساوي 72 ساعة وتحت الفرضيّات المذكورة آنفًا علماً أنَّ الانحراف المعياري $\sigma_o = ?$ مجهولاً و $n = 49$.

٤- حساب قيمة $\beta$ (احتمال الوقوع في الخطأ من النوع الثاني) إذا كان متوسط العمل الحقيقي للبطاريات يساوي 70 ساعة وبانحراف معياري $\sigma_o = 2$ ساعة، وكان:

أ- حجم العيّنة $n = 9$.

ب- حجم العيّنة $n = 36$.

ولنلاحظ ما الذي سيحدث لقيمة قوة الاختبار $1-\beta$؟

الحل: من أجل الطلب:

١- للرد على زعم المصنع علينا اختبار الفرضيّة الابتدائية $\mathbf{H}_o : \mu = \mu_o = 72$ وذلك عند مستوى الأهميّة $\alpha = 0.05$، وأمّا الفرضيّة البديلة فيمكننا اختيارها وفقاً لأحد النموذجين $\mathbf{H}_A : \mu \neq \mu_o = 72$ أو $\mathbf{H}_A : \mu < \mu_o = 72$، وذلك لأنّه من مُعطيات العيّنة لدينا متوسط أعمار هذه البطاريات في العيّنة يساوي $70 = \mu_o < 72$، وفي مثل هذه الدراسات يرجّح أخذ الفرضيّة البديلة $\mathbf{H}_A : \mu < \mu_o$ لأنَّ مُعطيات العيّنة تثير في نفسنا الشك تجاه حالة المساواة، وبالتالي من أجل هذه المسألة تكون نتيجة الاختبار المبنية على الاختيار $\mathbf{H}_A : \mu < \mu_o$ أكثر دقّةً.

الآن، وبما أنَّ $\sigma_o = 5$ فهذا يعني أنَّ الانحراف المعياري معلوماً، فإنَّه من أجل الاختيار $\mathbf{H}_o : \mu = \mu_o = 72$ لدينا من جدول التوزيع الطبيعي $z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$، وأمّا لحساب قيمة $z$ علينا استخدام العلاقة [11,7] حيث لدينا:

$$z = \frac{\bar{x} - \mu_o}{\sigma_o}\sqrt{n} = \frac{70-72}{5}\sqrt{36} = -2.40$$

ومن ثمَّ ينتج لدينا أنَّ $z = -2.40 < -1.96 = -z_{1-\frac{\alpha}{2}}$، وهكذا نجد أنَّ قيمة $z$ تقع في منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$، وبالتالي نرفض ادعاء المصنع حول متوسط زمن التشغيل للبطاريات عند مستوى الأهميّة $\alpha = 0.05$.

أمّا من أجل الفرضيّة البديلة $\mathbf{H}_A : \mu < \mu_o = 72$ نلاحظ أنَّ $z_{1-\alpha} = z_{0.95} = 1.645$، ومنه يكون لدينا:

$$z = -2.40 < -1.645 = -z_{1-\alpha}$$

ومن ثمَّ نحصل على القرار السابق نفسه أيضاً.

٢- من أجل الفرضيّة البديلة $\mathbf{H}_A : \mu \neq \mu_o = 72$ يكفي أن نعيّن قيمة $z_{1-\frac{\alpha}{2}}$ فنجدها تساوي $z_{1-\frac{\alpha}{2}} = z_{0.995} = 2.575$، ومنه يكون لدينا:

$$-z_{1-\frac{\alpha}{2}} = -2.575 < z = -2.40 < 2.575 = z_{1-\frac{\alpha}{2}}$$

وهكذا نجد أنَّ قيمة $z$ تقع في منطقة القبول للفرضيّة الابتدائية $\mathbf{H}_o$ ، وبالتالي نقبل ادعاء المصنع حول متوسط زمن تشغيل البطاريات عند مستوى الأهمّيّة $\alpha = 0.01$ .

أمّا لو كنّا قد أخذنا الفرضيّة البديلة $\mathbf{H}_A : \mu < \mu_o = 72$ فإنَّه سيكون $z_{1-\alpha} = z_{0.99} \approx 2.33$ ، ومنه يكون لدينا:

$$z = -2.40 < -2.33 = -z_{1-\alpha}$$

فنحصل من أجل هذه الحالة على قرار مختلف عن السابق حيث يكون قرارنا بالرفض في هذه الحالة عند مستوى الأهمّيّة $\alpha = 0.01$ .

٣- من أجل هذا الطلب علينا اختبار الفرضيّة الابتدائية $\mathbf{H}_o : \mu = \mu_o = 72$ عند مستوى الأهمّيّة $\alpha = 0.05$ ، وأمّا الفرضيّة البديلة فيمكننا اختيارها وفقاً لأحد لنموذجين $\mathbf{H}_A : \mu \neq \mu_o = 72$ أو $\mathbf{H}_A : \mu < \mu_o = 72$ ، ولكن في هذه الحالة لدينا الانحراف المعياري للمجتمع $\sigma^2$ مجهولاً و $n = 49$ ، ولذلك علينا حساب قيمة $t$ من خلال العلاقة [11,8] حيث لدينا:

$$t = \frac{\bar{x} - \mu_o}{s}\sqrt{n} = \frac{70-72}{4}\sqrt{49} = -3.50$$

ولذلك من أجل الفرضيّة البديلة $\mathbf{H}_A : \mu \neq \mu_o$ يجب حساب قيمة $t_{1-\frac{\alpha}{2}}^{[n-1]}$ من جدول توزيع ستودنت فنجد $t_{0.975}^{[48]} = 2.010$ ، ومن ثمَّ يكون لدينا:

$$t = -3.50 < -2.010 = -t_{0.975}^{[48]}$$

وهذا يعني أنَّ قيمة $t$ تقع في منطقة الرفض لـ $\mathbf{H}_o$ ، وبالتالي نرفض ادعاء المصنع حول متوسط زمن التشغيل للبطاريات عند مستوى الأهمّيّة $\alpha = 0.05$ .

أمّا من أجل الفرضيّة البديلة $\mathbf{H}_A : \mu < \mu_o$ فإنَّه يجب حساب قيمة $t_{1-\alpha}^{[n-1]}$ من جدول توزيع ستودنت فنجد $t_{0.99}^{[48]} = 2.407$ ، ومن ثمَّ يكون لدينا:

$$t = -3.50 < -2.407 = -t_{0.99}^{[48]}$$

وهذا يعني أنَّ قيمة $t$ تقع في منطقة الرفض للفرضيّة $\mathbf{H}_o$ أيضاً، ومن ثمَّ نرفض ادعاء المصنع حول متوسط زمن التشغيل للبطاريات أيضاً وذلك عند مستوى الأهمّيّة $\alpha = 0.05$ .

٤- لحساب قيمة $\beta$ نعلم أنَّ هذه القيمة هي احتمال قبول الفرضيّة الابتدائية $\mathbf{H}_o : \mu = \mu_o = 72$ علماً أنَّها خاطئة وذلك عند مستوى الأهمّيّة $\alpha$ ، ومن جهة أخرى، فإنَّ احتمال قبول الفرضيّة الابتدائية $\mathbf{H}_o : \mu = 70$ عند مستوى من الأهمّيّة $\alpha$ يساوي إلى المساحة التي تحت منحنى دالّة الكثافة الاحتمالية لـ $\bar{X}$ ، علماً أنَّ $\bar{X} \overset{\infty}{\sim} \mathbf{N}\left(\mu, \frac{\sigma_o}{\sqrt{n}}\right)$ فوق فترة الثقة الآتية:

$$\mathbf{CI} = \left(\mu_o - z_{1-\frac{\alpha}{2}}\frac{\sigma_o}{\sqrt{n}} \quad , \quad \mu_o + z_{1-\frac{\alpha}{2}}\frac{\sigma_o}{\sqrt{n}}\right)$$

أ- فإذا كان حجم العيّنة $n = 9$ ، فإنَّه سيكون لفترة الثقة السابقة العرض الآتي:

$$\left(72 - 1.96\frac{2}{\sqrt{9}} \; , \; 72 + 1.96\frac{2}{\sqrt{9}}\right) = (70.693 \; , \; 73.307)$$

ومن ثمَّ يكون لدينا:

$$\beta = P\left(70.693 < \bar{X} < 73.307 \,|\, H_o : \mu = 72 \text{ is false}\right) = P\left(70.693 < \bar{X} < 73.307 \,|\, H_o : \mu = 70 \text{ is true}\right)$$
$$= P\left(\frac{70.693-70}{2/3} < \frac{\bar{X}-70}{2/3} < \frac{73.307-70}{2/3}\right)$$

والآن بوضع $Z = \dfrac{\bar{X}-70}{2/3}$ ينتج لدينا الآتي:

$$\beta = P\left(1.04 < Z < 4.96\right) = P\left(Z < 4.96\right) - P\left(Z \le 1.04\right) \approx 1 - 0.8508 = 0.1492$$

ومن ثَّم تكون قيمة قوة الاختبار هي $1-\beta \approx 1 - 0.1492 = 0.8508$ .

ب- أمّا إذا كان حجم العيّنة $n = 36$ ، فإنّه سيكون لفترة الثقة السابقة العرض الآتي:

$$\left(72 - 1.96\frac{2}{\sqrt{36}} \,,\, 72 + 1.96\frac{2}{\sqrt{36}}\right) = (71.347 \,,\, 72.953)$$

وبالتالي يكون لدينا:

$$\beta = P\left(71.347 < \bar{X} < 72.953 \,|\, H_o : \mu = 72 \text{ is false}\right) = P\left(71.347 < \bar{X} < 72.953 \,|\, H_o : \mu = 70 \text{ is true}\right)$$
$$= P\left(\frac{71.347-70}{2/6} < \frac{\bar{X}-70}{2/6} < \frac{72.953-70}{2/6}\right)$$

وبوضع $Z = \dfrac{\bar{X}-70}{2/6}$ ينتج لدينا الآتي:

$$\beta = P\left(4.041 < Z < 8.859\right) = P\left(Z < 8.859\right) - P\left(Z \le 4.041\right) \approx 1 - 1 = 0$$

ومن ثَّم تكون قيمة قوة الاختبار مساوية لـ $1-\beta \approx 1$ ، وهكذا نلاحظ أنَّ ازدياد حجم العيّنة أدى إلى تصغير قيمة $\beta$ بشكل كبير ومن ثَّم تحسين قيمة قوة الاختبار بشكل ملحوظ.

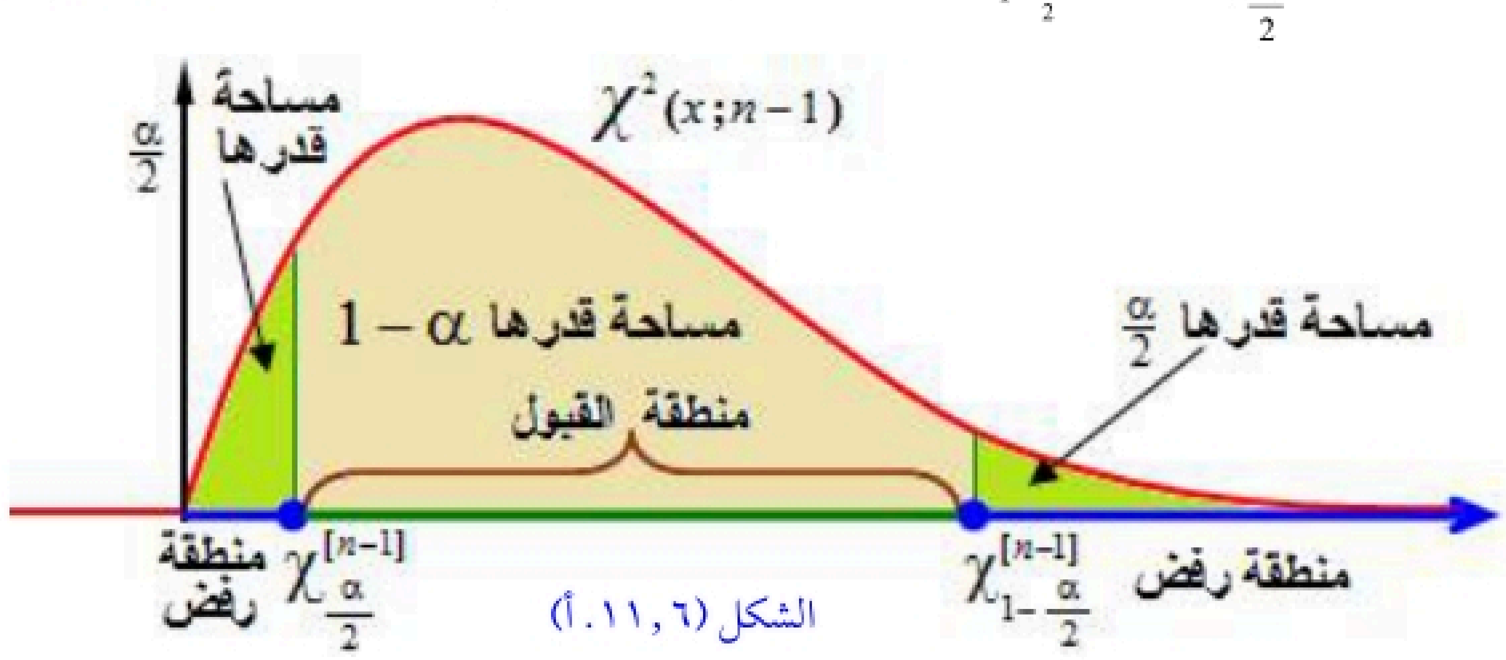
### (١١,٢,٢) اختبارات من أجل تباين مجتمع طبيعي

تُقدِّم لنا المبرهنة الآتية (سنقبلها دون برهان) اختباراً لفرضيات متعلّقة بتباين مجتمع طبيعي.

#### (١١,٢,٢,١) مبرهنة

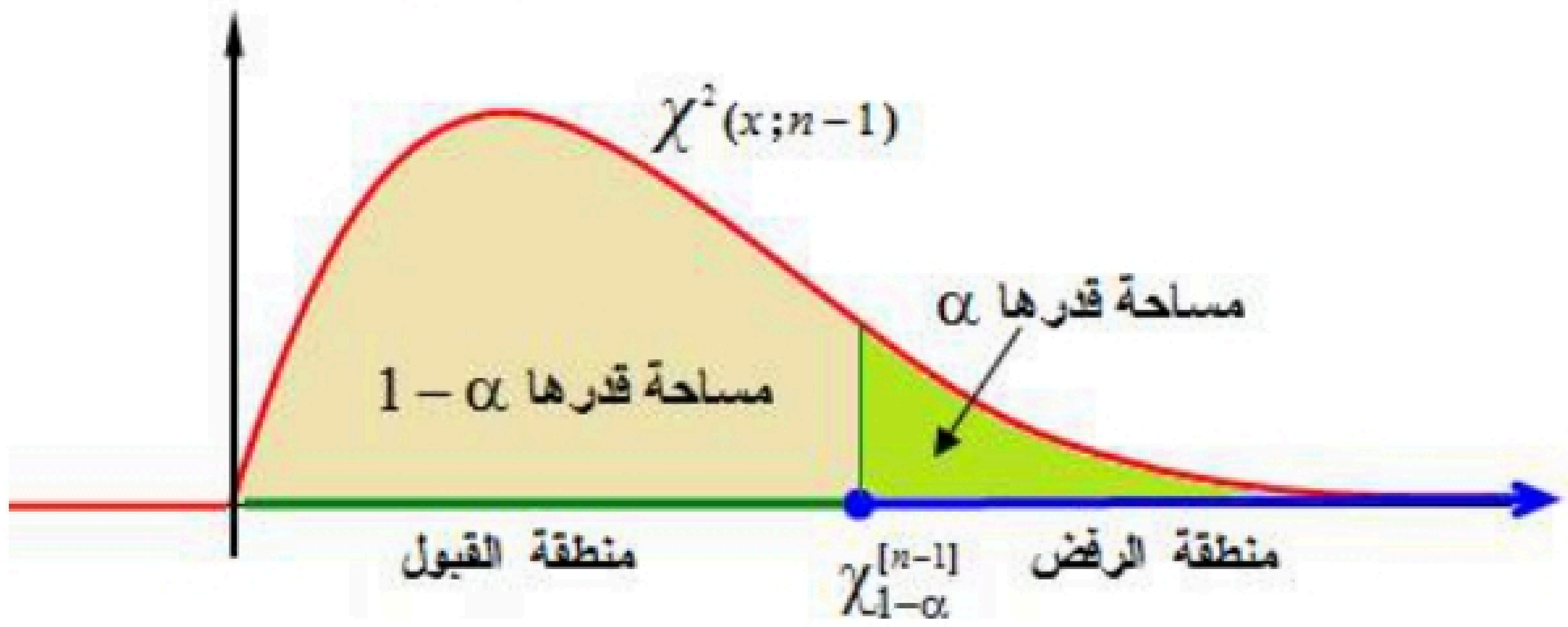
ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً إحصائياً موصوفاً من خلال متغيّر عشوائي $X$ خاضع للتوزيع $\mathbf{N}(\mu, \sigma)$ ولنأخذ $\mathcal{X}_n$ عيّنة من هذا المجتمع قيمة متوسطها $\bar{x}$ وانحرافها المعياري $s$ ، ولنفترض أنَّ قيمة متوسط هذا المجتمع $\mu = \mu_o$ معلوماً. عندئذ من أجل مستوى من الأهمّيّة قدره $\alpha$ تكون منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o : \sigma^2 = \sigma_o^2$ مُعيّنة بوساطة:

١- المتباينتين $\chi \le \chi^{[n-1]}_{\frac{\alpha}{2}}$ و $\chi \ge \chi^{[n-1]}_{1-\frac{\alpha}{2}}$ إذا كانت الفرضيّة البديلة $\mathbf{H}_A : \sigma^2 \ne \sigma_o^2$ (انظر الشكل الآتي).



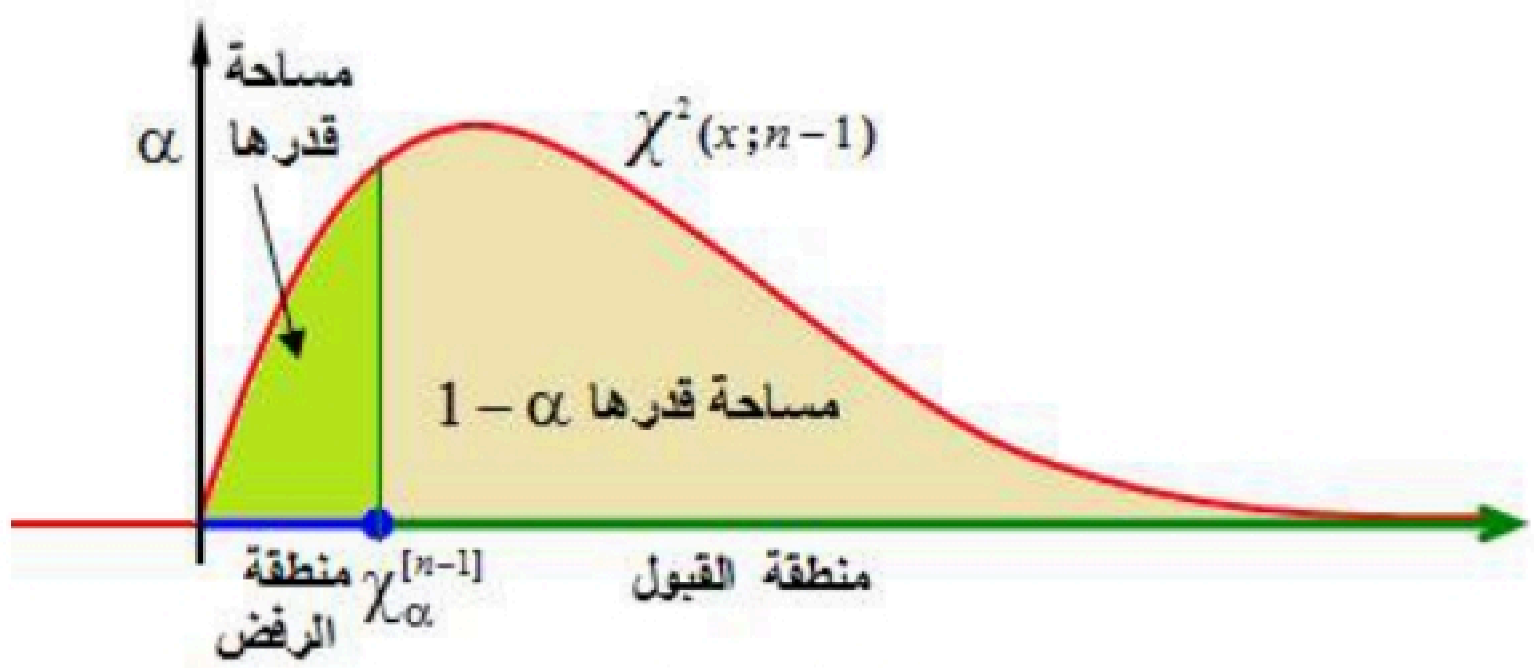

الشكل (١١,٦.أ)

٢- المتباينة $\chi \le \chi_{1-\alpha}^{[n-1]}$ إذا كانت الفرضيّة البديلة $\mathbf{H}_A : \sigma^2 > \sigma_o^2$ (انظر الشكل الآتي).




الشكل (٦ , ١١ . ب)

٣- المتباينة $\chi \ge \chi_{\alpha}^{[n-1]}$ إذا كانت الفرضيّة البديلة $\mathbf{H}_A : \sigma^2 < \sigma_o^2$ (انظر الشكل الآتي).




الشكل (٦ , ١١ . ج)

علماً أنَّ القيمة $\chi$ تُحسب من العلاقة الآتية:

$$\chi = \frac{(n-1)\, s^2}{\sigma_o^2} \qquad [11,9]$$

**(٢ , ٢ , ٢ , ١١) مثال**

زعمت شركة لإنتاج إطارات السيّارات أنّها على ثقة قدرها 0.99 بأنَّ الإطارات المصنَّعة لديها تتحمّل في المتوسط مسير 50000 كيلو متر بانحراف معياري قدره 10000 كيلو متر. الآن ومن أجل التّحقُّق من زعم هذه الشركة قمنا بسحب عيّنة من إنتاج هذه الشركة بحجم $n = 36$ إطاراً فوجدنا أنَّ إطارات هذه العيّنة تتحمّل في المتوسط مسير 48000 كيلو متر بانحراف معياري قدره 11000 كيلو متر، فإذا افترضنا أنَّ القياسات الناتجة المسافات التي تتحملها هذه الإطارات تتبع التوزيع الطبيعي، فعندئذ:

١- ما هو ردنا على ادّعاء هذه الشركة بخصوص متوسط المسافة التي تتحمَّلها الإطارات المصنَّعة لديه إذا أخذنا بقيمة الانحراف المعياري المقدَّم من قبلها؟

٢- ما هو ردنا على ادّعاء هذه الشركة بخصوص متوسط المسافة التي تتحمَّلها الإطارات المصنَّعة لديه إذا تجاهلنا قيمة الانحراف المعياري المقدَّم من قبلها، وذلك لأنَّ مُعطيات العيّنة أثارت لدينا الشك حيال قيمة الانحراف المعياري المقدَّم ؟

٣- ما هو ردنا على ادّعاء هذه الشركة بخصوص قيمة التباين المقدَّم من قبلها حول المسافة التي تتحمَّلها الإطارات المصنَّعة لديها؟

الحـل: من أجل:

١- من أجل هذا الطلب علينا أن نختبر الفرضيّة الابتدائية $\mathbf{H}_o : \mu = \mu_o = 50000$ مقابل الفرضيّة البديلة $\mathbf{H}_A : \mu < \mu_o = 50000$ عند مستوى من الأهمّيّة $\alpha = 0.01$ لأنَّ مستوى الثقة المقدَّم من قبل الشركة يساوي $1-\alpha = 0.99$.

بالطبع اخترنا هذا النموذج من الفرضيّات البديلة، لأنّنا لاحظنا الفرق الكبير بين القيمة المقدَّمة من قبل الشركة وتلك التي حصلنا عليها من العيّنة. الآن وبما أنّنا سنأخذ بقيمة الانحراف المعياري المقدَّم من قبل الشركة فإنَّ ذلك يعني أنَّ تباين مجتمع الإطارات معلوم ويساوي $\sigma_o^2 = 10^8$ كيلو متر، ومن ثَّم تكون منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ هي المنطقة المعيَّنة بالمتباينة $z \leq -z_{1-\alpha}$، علماً أنَّ قيمة $z$ تحسب بالعلاقة [11,7]، حيث لدينا:

$$z = \frac{\bar{x} - \mu_o}{\sigma_o}\sqrt{n} = \frac{48000 - 50000}{10000}\sqrt{36} = -1.20$$

ومن جهة أخرى نجد $z_{1-\alpha} = z_{0.99} = 2.33$ (اخترنا هذه القيمة لأنّه لدينا اختبار وحيد الطرف)، وبملاحظة أنَّ:

$$z = -1.20 > -2.33 = -z_{1-\alpha}$$

فإنَّ ذلك يعني أنَّ قيمة $z$ تقع في منطقة القبول للفرضيّة الابتدائية، ومن ثَّم نقبل بادعاء الشركة حول المسافة التي تتحمَّلها إطاراتها، وذلك عند مستوى الأهمّيّة $\alpha = 0.01$.

لاحظ هنا حتى لو أخذنا الفرضيّة البديلة $\mathbf{H}_A : \mu \neq \mu_o = 48000$ فإنّنا سنحصل على قبول ادعاء الشركة أيضاً (بخصوص المسافة التي تتحمَّلها الإطارات) عند مستوى الأهمّيّة $\alpha = 0.01$، وذلك لأنّه لدينا $z_{1-\frac{\alpha}{2}} = z_{0.995} = \frac{2.57+2.58}{2} = 2.575$، فينتج لدينا من أجل هذه الحالة:

$$z = -1.20 > -2.575 = -z_{1-\frac{\alpha}{2}}$$

٢- في هذه الحالة لدينا تباين مجتمع الإطارات $\sigma^2$ مجهولاً، ولذلك للرد على ادعاء الشركة علينا أنَّ نختبر الفرضيّة الابتدائية $\mathbf{H}_o : \mu = \mu_o = 50000$ مقابل الفرضيّة البديلة $\mathbf{H}_A : \mu = 48000 < \mu_o = 50000$ وذلك عند مستوى من الأهمّيّة قدره $\alpha = 0.01$، والذي دفعنا إلى أخذ هذه الفرضيّة البديلة هو ذات السبب الذي ذكرناه في الطلب السابق.

الآن، ولتنفيذ هذا الاختبار علينا حساب قيمة $t$ من خـلال العلاقة [11,8]، فنجد ما يلي:

$$t = \frac{\bar{x} - \mu_o}{s}\sqrt{n} = \frac{48000 - 50000}{11000}\sqrt{36} = -1.091$$

من جهة أخرى لدينا $t_{1-\alpha}^{[n-1]} = t_{0.99}^{[35]} = 2.438$، ومنه ينتج لدينا أنَّ:

$$t = -1.091 > -2.438 = -t_{0.99}^{[35]}$$

ومن ثَّم نقبل في هذه الحالة بادعاء الشركة أيضاً (حول المسافة التي تتحمَّلها إطاراتها)، وذلك عند مستوى الأهمّيّة $\alpha = 0.01$.

٣- من أجل هذا الطلب علينا اختبار الفرضيّة الابتدائية $\mathbf{H}_o : \sigma^2 = \sigma_o^2 = 10^8$ مقابل الفرضيّة البديلة: $\mathbf{H}_A : \sigma^2 > \sigma_o^2 = 10^8$ عند مستوى من الأهمّيّة قدره $\alpha = 0.01$، وقد اخترنا هذا النموذج من الفرضيّات البديلة لأنّنا لاحظنا الفرق الكبير بين القيمة المقدَّمة من قبل الشركة وتلك التي حصلنا عليها من العيّنة حيث تولَّد لدينا شك حيال قيمة التباين المقدَّمة من قبل الشركة. الآن، ولتنفيذ هذا الاختبار علينا حســاب قيمة الإحصاءة $\chi$ من خلال العلاقة [11,9]، حيث لدينا:

$$\chi = \frac{(n-1)s^2}{\sigma_o^2} = \frac{35\ (11000)^2}{10^8} = 42.35$$

وكذلك علينا تعيين قيمة $\chi_{1-\alpha}^{[n-1]}$ من جدول توزيع كاي مربّع فنجدها تساوي $\chi_{1-\alpha}^{[n-1]} = \chi_{0.99}^{[35]} = 57.342$ ، ومن ثمَّ يكون لدينا:

$$\chi_{1-\alpha}^{[n-1]} = 57.342 > 42.35 = \chi$$

وهذا يعني أنَّ قيمة $\chi$ تقع في منطقة القبول للفرضيّة الابتدائية $\mathbf{H}_o$، ومن ثمَّ ليس أمامنا إلاَّ أن نقبل بادعاء الشركة حول التباين المقدَّم من قبلها، وذلك عند مستوى الأهميّة $\alpha = 0.01$ .

لاحظ هنا حتى لو أخذنا مقابل الفرضيّة البديلة $\mathbf{H}_A : \sigma^2 \neq \sigma_o^2 = 10^8$ فإنَّنا سنحصل على قبول ادعاء الشركة (بخصوص التباين) أيضاً، وذلك عند مستوى الأهميّة $\alpha = 0.01$ ، وذلك لأنَّه لدينا $\chi_{1-\frac{\alpha}{2}}^{[n-1]} = \chi_{0.995}^{[35]} = 60.275$ وكذلك $\chi_{\frac{\alpha}{2}}^{[n-1]} = \chi_{0.005}^{[35]} = 17.192$ حيث نلاحظ أنَّ قيمة $\chi = 42.35$ تقع في منطقة القبول للفرضيّة الابتدائية $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha = 0.01$ .

## (١١,٣) الاختبارات الإحصائية من أجل معلمة مجتمع برنولي

**Statistical Tests for a Parameter of Bernoulli Population**

المبرهنة الآتية (سنقبلها دون برهان أيضاً) تُقدِّم لنا اختباراً إحصائياً لفرضيّة تتعلّق بمَعْلَمة مجتمع برنولي.

### (١١,٣,١) مبرهنة

ليكن $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع لتوزيع برنولي بمَعْلَمة $p$ (احتمال نجاح)، ولنأخذ $\mathcal{X}_n$ عيّنة من هذا المجتمع، ولتكن لدينا الفرضيّة الابتدائية $\mathbf{H}_o : p = p_o$ ، فعندئذ من أجل مستوى للأهميّة قدره $\alpha$ تكون:

١ –منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ مُعيَّنةً بالفترة المحدَّدة بالمتباينة $|z| \geq z_{1-\frac{\alpha}{2}}$ (انظر الشكل السابق (١١,٤ . أ)) إذا كانت الفرضيّة البديلة من النموذج $\mathbf{H}_A : p \neq p_o$ ، علماً أنَّ $z$ تُحسب بالعلاقة الآتية:

$$z = \frac{\bar{p} - p_o}{\sqrt{p_o\,(1-p_o)}}\sqrt{n} \qquad [11,10]$$

وأمّا $\bar{p}$ فهي قيمة الإحصاءة $\overline{\pi}$ الذي يمثِّل التكرار النسبي للنجاحات في العيّنة $\mathcal{X}_n$ .

٢– منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ مُعيَّنةً بالفترة المحدَّدة بالمتباينة $z \leq -z_{1-\alpha}$ (انظر الشكل السابق (١١,٤ . ب)) إذا كانت الفرضيّة البديلة من النموذج $\mathbf{H}_A : p < p_o$ ، علماً أنَّ $z$ تُحسب بالعلاقة [11,10].

٣– منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ مُعيَّنةً بالفترة المحدَّدة بالمتباينة $z \geq z_{1-\alpha}$ (انظر الشكل السابق (١١,٤ . ج)) إذا كانت الفرضيّة البديلة من النموذج $\mathbf{H}_A : p > p_o$ ، علماً أنَّ $z$ تُحسب بالعلاقة [11,10].

### (١١,٣,٢) مثال

ادعت إحدى المدارس الشهيرة التي تحوي عدداً كبيراً جداً من الطلاب أنّها على ثقة قدرها 98% بأنَّ 15% من طلابها (في مختلف الفصول والسنوات) على الأقل سيحصلون على التقدير $A^+$ في الاختبار النهائي لمقرّر الرياضيات، ومن أجل التّحقُّق من صحة ادعاء هذه المدرسة قمنا بسحب عيّنة حجمها $n = 81$ طالباً من الذين تقدموا للاختبار النهائي لمقرّر الرياضيات، فوجدنا من بينهم 11 طالباً قد حصلوا على التقدير $A^+$ ، والسؤال: **هل تؤيد مُعطيات هذه العيّنة صحة ادّعاء المدرسة؟**

ومن ثمَّ لنبحث عن الإجابة على هذا السؤال إذا كنّا قد أعدنا سحب العيّنة بحيث تشتمل 450 طالباً، وأنّنا قد وجدنا من بينهم 50 طالباً قد حصلوا على التقدير $A^+$.

الحـل: نلاحظ هنا أنَّ الكشف عن نتيجة كل طالب تقدَّم للاختبار النهائي لمقرّر الرياضيات هي تجربة برنولية، فإمّا أن يكون قد حقَّق النتيجة $A^+$ أو لم يتمكن من ذلك (نجاح إذا حقَّق النتيجة +A، أو فشل إذا لم يحقِّق النتيجة +A)، وهذه النتائج مستقل بعضها عن بعض (تحت الافتراض بأنَّ الاختبار كان نزيهاً)، ولذلك حتى نجيب على هذا السؤال المطروح علينا اختبار الفرضيّة الابتدائية $\mathbf{H}_o : p = p_o = 0.15$ مقابل الفرضيّة البديلة $\mathbf{H}_A : p \neq p_o = 0.15$ عند مستوى الأهميّة $\alpha = 0.02$، وذلك لأنَّ مستوى الثقة التي قدَّمته المدرسة يساوي $1-\alpha = 0.98$.

الآن، ومن أجل تنفيذ هذا الاختبار يجب حساب قيمة $\bar{p}$ ومن ثمَّ $z$، حيث لدينا $\bar{p} = \frac{11}{81} = 0.1358$، وأمّا قيمة $z$ فإنّها تحسب وفقاً للعلاقة [11,10]، فنجدها تساوي:

$$z = \frac{\bar{p} - p_o}{\sqrt{p_o\,(1-p_o\,)}}\sqrt{n} = -0.358$$

وأخيراً علينا تعيين قيمة $\mathfrak{z}_{1-\frac{\alpha}{2}} = \mathfrak{z}_{0.99}$ من جدول التوزيع الطبيعي المعياري، فنجدها تساوي $\mathfrak{z}_{1-\frac{\alpha}{2}} = \mathfrak{z}_{0.99} \approx 2.33$، وهكذا يكون لدينا:

$$|z\,| = 0.358 < 2.33 \approx \mathfrak{z}_{0.99}$$

ومن ثمَّ قيمة $z$ تقع في منطقة القبول للفرضيّة الابتدائية $\mathbf{H}_o$، ومن ثمّ ليس أمامنا إلاَّ القبول بادعاء المدرسة عند مستوى الأهميّة $\alpha = 0.02$.

الآن، ومن أجل المعطيات الجديدة للعينة الثانية نجد أنَّ $\bar{p} = \frac{50}{450} = 0.11$، وأمّا قيمة $z$ فتحسب وفقاً للعلاقة [11,10]، فنجدها تساوي:

$$z = \frac{\bar{p} - p_o}{\sqrt{p_o\,(1-p_o\,)}}\sqrt{n} = \frac{0.11-0.15}{\sqrt{0.15\,(1-0.15\,)}}\sqrt{450} = -2.376$$

وهكذا يكون لدينا:

$$\mathfrak{z}_{0.99} \approx 2.33 < 2.376 = |z\,|$$

ومن ثمَّ قيمة $z$ تقع في منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$، وبناءً على ذلك نرفض ادعاء المدرسة عـند مستوى الأهميّة $\alpha = 0.02$.

نلاحظ أنّه في الطلب السّابق لم يكن لدينا من مُعطيات العيّنة ما يدفعنا لرفض الفرضيّة الابتدائية $\mathbf{H}_o$، في حين أنَّه بزيادة حجم العيّنة (ومن ثمَّ زيادة كميّة المعلومات حول المَعْلَمة $p$) حصلنا على رفض لهذه الفرضيّة وهو يُعدُّ إقراراً بخطئها.

## (١١,٤) اختبارات من أجل الفرق بين متوسطي مجتمعين طبيعيين مستقلين

تُقدِّم لنا المبرهنة الآتية (سنقبلها جميعاً دون برهان) اختبارات لفرضيات متعلِّقة بالفرق بين متوسطي مجتمعين طبيعيين.

### (١١,٤,١) مبرهنة

ليكن لدينا $[\Omega_1\,,\,\mathcal{S}_1\,,\,\mathcal{P}_1\,]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع للتوزيع $\mathbf{N}(\mu_1\,,\sigma_1)$ ولنأخذ $\mathcal{X}_n$ عيّنة من هذا المجتمع. كذلك ليكن $[\Omega_2\,,\,\mathcal{S}_2\,,\,\mathcal{P}_2\,]$ مجتمعاً آخر مستقل عن $[\Omega_1\,,\,\mathcal{S}_1\,,\,\mathcal{P}_1\,]$ وموصوفاً من خلال متغيِّر عشوائي $Y$ خاضع

للتوزيع $N(\mu_2, \sigma_2)$، ولنأخذ $\mathcal{Y}_m$ عيّنة مسحوبة من هذا المجتمع الأخير، فإذا كانت $H_o: \mu_1 - \mu_2 = a_o$ فرضية ابتدائية يُطلب اختبارها عند مستوى من الأهمية $\alpha$، علماً أنَّ $a_o$ ثابت حقيقي، فعندئذ منطقة الرفض لهذه الفرضيّة تُعيَّن على النحو الآتي:

١- إذا كان الانحراف المعيّاري (أو التباين) للمجتمعين معلوماً ولهما القيم $\sigma_1 = \sigma_{1,o}$ و $\sigma_2 = \sigma_{2,o}$ فعندئذ:

أ- تُعطى منطقة الرفض للفرضيّة الابتدائية $H_o$ بوساطة المتباينة $|z| \geq z_{1-\frac{\alpha}{2}}$ إذا كانت الفرضيّة البديلة من النموذج:

$H_A: \mu_1 - \mu_2 \neq a_o$

ب- تُعطى منطقة الرفض للفرضيّة الابتدائية $H_o$ بوساطة المتباينة $z \geq z_{1-\alpha}$ إذا كانت الفرضيّة البديلة من النموذج:

$H_A: \mu_1 - \mu_2 > a_o$

ج- تُعطى منطقة الرفض للفرضيّة الابتدائية $H_o$ بوساطة المتباينة $z \leq -z_{1-\alpha}$ إذا كانت الفرضيّة البديلة من النموذج:

$H_A: \mu_1 - \mu_2 < a_o$

علماً أنَّ $z$ تُحسب بوساطة العلاقة الآتية:

$$z = \frac{(\bar{x} - \bar{y}) - a_o}{\sqrt{m\,\sigma_{1,o} + n\,\sigma_{2,o}}} \sqrt{n \cdot m} \qquad [11,11]$$

٢- إذا كان الانحراف المعياري (أو التباين) للمجتمعين مجهولاً $\sigma_1 = ?$ و $\sigma_2 = ?$، ولكن مع الافتراض بتساويهما (أي إنَّ $\sigma_1 = \sigma_2$)، فعندئذ:

أ- تُعطى منطقة الرفض للفرضيّة الابتدائية $H_o$ بوساطة المتباينة $|t| \geq t_{1-\frac{\alpha}{2}}^{[n+m-2]}$ إذا كانت الفرضيّة البديلة من النموذج:

$H_A: \mu_1 - \mu_2 \neq a_o$

ب- تُعطى منطقة الرفض للفرضيّة الابتدائية $H_o$ بوساطة المتباينة $t \geq t_{1-\alpha}^{[n+m-2]}$ إذا كانت الفرضيّة البديلة من النموذج:

$H_A: \mu_1 - \mu_2 > a_o$

ج- تُعطى منطقة الرفض للفرضيّة الابتدائية $H_o$ بوساطة المتباينة $t \leq t_{1-\alpha}^{[n+m-2]}$ إذا كانت الفرضيّة البديلة من النموذج:

$H_A: \mu_1 - \mu_2 < a_o$

علماً أنَّ قيمة $t$ تُحسب بوساطة العلاقة الآتية:

$$t = \frac{\left[(\bar{x} - \bar{y}) - a_o\right] \sqrt{n \cdot m \cdot (n+m-2)}}{\sqrt{(m+n)\left[(n-1)S_X^2 + (m-1)S_Y^2\right]}} \qquad [11,12]$$

### (١١,٤,٢) ملاحظات

١- في الحالة الخاصة عندما يكون لدينا $a_o = 0$ فإنَّ الاختبارات السابقة في المبرهنة (١١-٤-١) تؤول إلى اختبارات متعلّقة بتساوي متوسطي مجتمعين طبيعيين مستقلّين.

٢- في حال كان $\sigma_1$ و $\sigma_2$ مجهولين ولم تُقدَّم أيّة معلومات حول تساويهما، فعندئذ يجب التّحقُّق أولاً إن كان $\sigma_1^2 = \sigma_2^2$ باحتمال قدره $1-\alpha$ أم لا، وذلك من خلال تعيين $100(1-\alpha)\%$ فترة ثقة للنسبة $\sigma_1^2 / \sigma_2^2$ والمُعطاة بالفترة [10-34-b]، فإذا كان العدد (1) ينتمي إلى هذه الفترة، فعندئذ يكون لدينا $\sigma_1^2 = \sigma_2^2$ باحتمال قدره $1-\alpha$، ومن ثمَّ نتابع بتطبيق الاختبار الخاص بالحالة $\sigma_1^2 = \sigma_2^2$. أمّا إذا كان العدد (1) لا ينتمي إلى تلك الفترة، فعندئذ نتوقف عن تنفيذ الاختبار لأنّه لا يُحقِّق الفرضيّات المطلوبة، والاختبار الذي يعالج مثل هذه الحالة لن نتطرق إليه.

## (١١,٥) اختبارات من أجل نسبة التباين لمجتمعين طبيعيين مستقلين

تُقدِّم لنا المبرهنة الآتية اختباراً لفرضيّة تتعلّق بنسبة التباين لمجتمعين طبيعيين مستقلَّين.

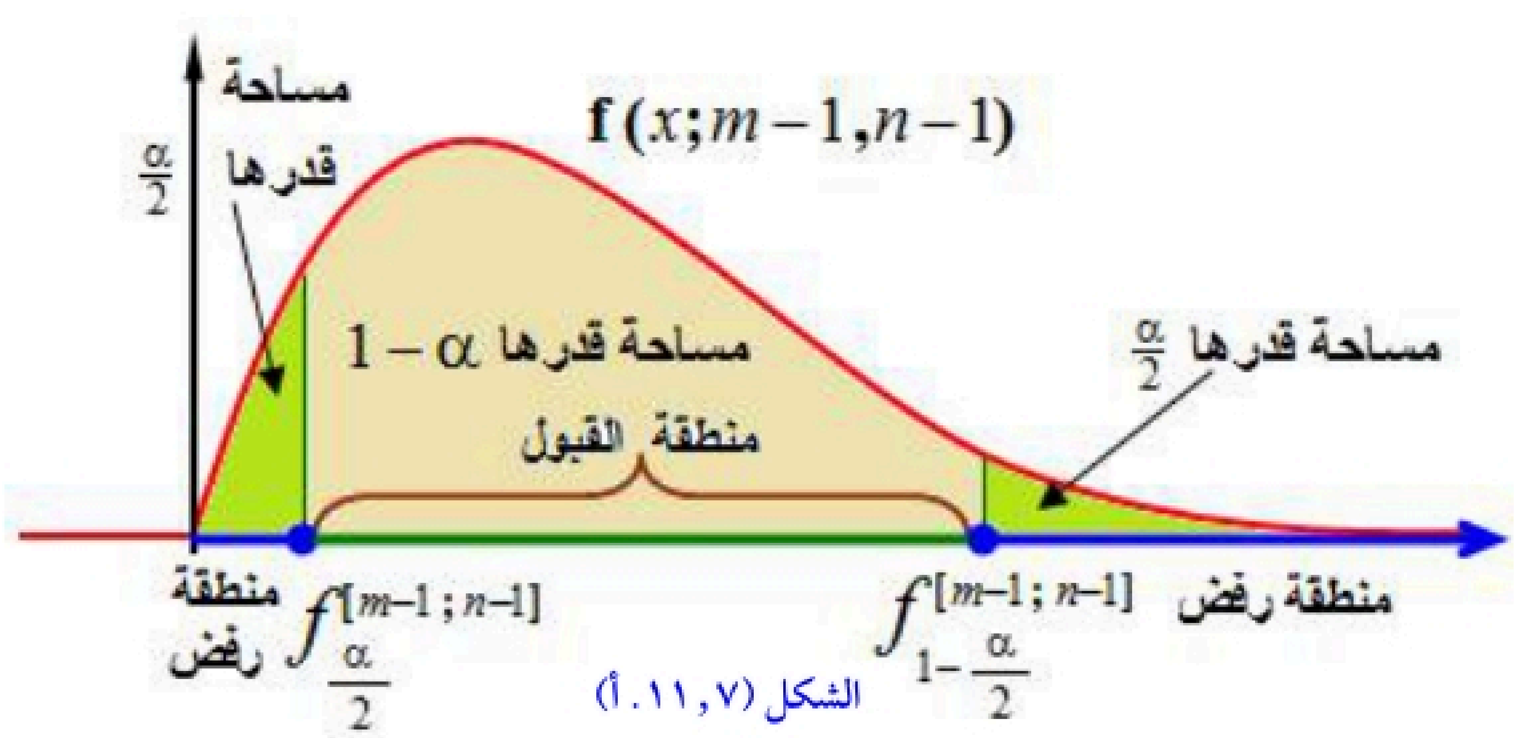
### (١١,٥,١) مبرهنة

ليكن لدينا $[\Omega_1, \mathbb{S}_1, \mathcal{P}_1]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ خاضع للتوزيع $\mathbf{N}(\mu_1, \sigma_1)$ ولنأخذ $\mathcal{X}_n$ عيّنة من هذا المجتمع. كذلك ليكن $[\Omega_2, \mathbb{S}_2, \mathcal{P}_2]$ مجتمعاً آخر مستقل عن المجتمع $[\Omega_1, \mathbb{S}_1, \mathcal{P}_1]$ وموصوفاً من خلال متغيِّر عشوائي $Y$ خاضع للتوزيع $\mathbf{N}(\mu_2, \sigma_2)$، ولنأخذ $\mathcal{Y}_m$ عيّنة مسحوبة من المجتمع الثاني. الآن بفرض أنّ $\mathbf{H}_o : \frac{\sigma_1^2}{\sigma_2^2} = u_o$ فرضيّة ابتدائية يُطلب اختبارها عند مستوى من الأهميّة $\alpha$، علماً أنّ $u_o$ ثابت حقيقي، فعندئذ منطقة الرفض لهذه الفرضيّة تُعيَّن على النحو الآتي:

١- تُعطى منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ بوساطة المتباينتين الآتيتين:

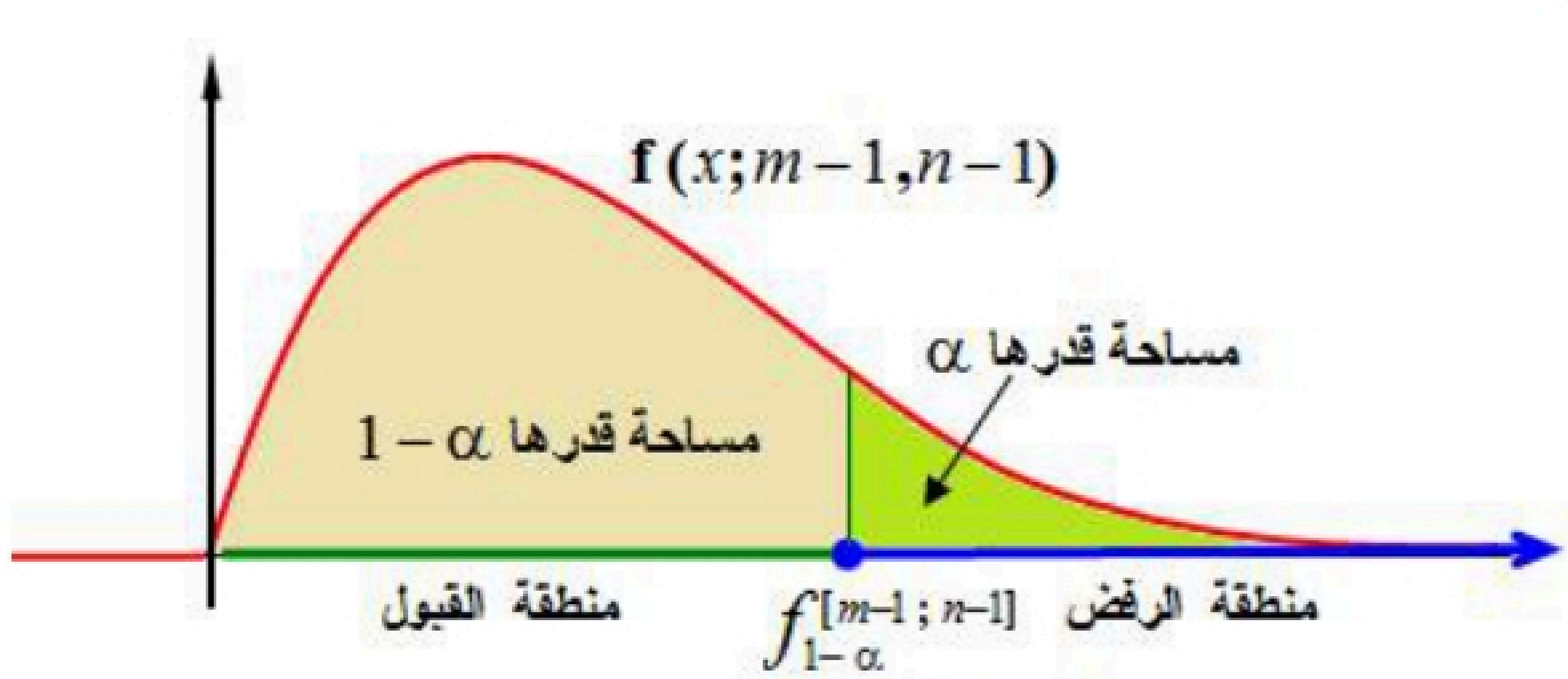
$$f \le f_{\frac{\alpha}{2}}^{[m-1\,;\,n-1]} = \left( f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]} \right)^{-1} \qquad \& \qquad f \ge f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]}$$

إذا كانت الفرضيّة البديلة من النموذج $\mathbf{H}_A : \frac{\sigma_1^2}{\sigma_2^2} \ne u_o$ (للتوضيح انظر الشكل الآتي (١١,٧ أ)).



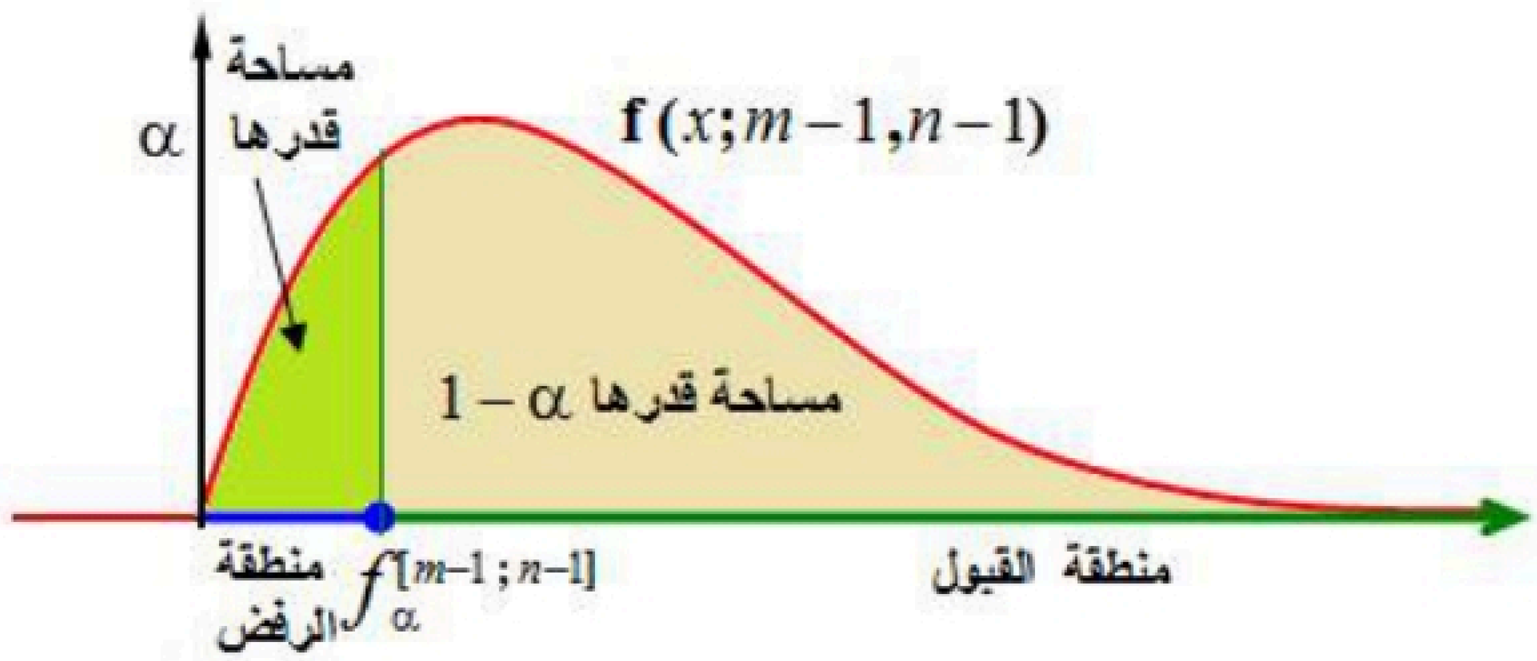

الشكل (١١,٧ أ)

٢- تُعطى منطقة الرفض للفرضيّة $\mathbf{H}_o$ بوساطة المتباينة $f_{1-\alpha}^{[m-1\,;\,n-1]} \le f$ إذا كانت الفرضيّة البديلة $\mathbf{H}_A : \frac{\sigma_1^2}{\sigma_2^2} > u_o$ (للتوضيح انظر الشكل الآتي (١١,٧ ب)).




الشكل (١١,٧ ب)

٣- تُعطى منطقة الرفض للفرضية $\mathbf{H}_o$ بوساطة المتباينة $f \geq f_{\alpha}^{[m-1\,;\,n-1]}$ إذا كانت الفرضيّة البديلة $\mathbf{H}_o : \frac{\sigma_1^2}{\sigma_2^2} < u_o$ (للتوضيح انظر الشكل الآتي (٧ ,١١ أ)).




الشكل (٧ ,١١ .ج)

علماً أنَّ $f_{\alpha}^{[m-1\,;\,n-1]} = \left(f_{1-\alpha}^{[m-1\,;\,n-1]}\right)^{-1}$ وأمّا قيمة $f$ فإنَّها تُحسب بوساطة العلاقة الآتية:

$$f = \frac{s_2^2}{s_1^2} \cdot u_o \qquad [11,13]$$

### (٢, ٥, ١١) ملاحظة

في الحالة الخاصة عندما يكون $u_o = 1$ فإنَّ الاختبارات السابقة في المبرهنة (١ ,٥ ,١١) تؤول إلى اختبارات متعلِّقة بتساوي التباين لمجتمعين طبيعيين مستقلَّين.

### (٣, ٥, ١١) مثال

تقدَّمت شركتان $A_1$ و $A_2$ لإنتاج الدواء بعروضهما إلى وزارة الصحة لتزويدها بنوع من أدوية الالتهابات (وليكن على سبيل المثال كبسولات أموكسيل عيار **500mg**) حيث ادعت الشركة $A_1$ أنَّها على ثقة قدرها 99% بأنَّ متوسط وزن المادة الفعَّالة في منتجها المقدَّم 500 mg بانحراف معياري قدره 25mg، في حين ادعت الشركة الأخرى $A_2$ أنَّها على ثقة قدرها 99% أيضاً بأنَّ متوسط وزن المادة الفعَّالة في منتجها المقدَّم 500mg بانحراف معياري قدره 25mg أيضاً. الآن لكي تحدِّد الوزارة أي الشركتين ستختار قامت بسحب عيّنة $\mathcal{X}_{36}$ من منتج الشركة $A_1$ وأخضعتها للتحليل فوجدت أنَّ متوسط وزن المادة الفعَّالة في هذه العيّنة 480mg بانحراف معياري قدره 20mg، وكذلك قامت بسحب عيّنة أخرى $\mathcal{Y}_{49}$ من منتج الشركة $A_2$ وأخضعتها للتحليل فوجدت أنَّ متوسط وزن المادة الفعَّالة في هذه العيّنة 505mg بانحراف معياري قدره 15mg . الآن، وتحت الفرض أنَّ أوزان المادَّة الفعَّالة في الدواء له توزيع طبيعي، لنقم بما يلي:

١- تحديد أي الشركتين يجب اختيارها لتوريد الدواء إذا كانت الوزارة:

أ- ستأخذ بقيم الانحراف المعياري المقدَّم من قبل الشركتين؟

ب- ستتجاهل قيم الانحراف المعياري المقدَّم من قبل الشركتين؟

٢- لنختبر الفرضيّة القائلة بتساوي التباين لدى مجتمعي أدوية الشركتين.

الحـل: سنفترض أنَّ $X$ و $Y$ هو المتغيِّر العشوائي الواصف لمجتمع كبسولات الشركة $A_1$ و $A_2$ على الترتيب، فيكون لدينا $X \sim \mathbf{N}(\mu_1, \sigma_1)$ و $Y \sim \mathbf{N}(\mu_2, \sigma_2)$، ومن ثمَّ لكي نحدِّد أي الشركتين يجب اختيارها لتوريد الدواء إلى الوزارة علينا اختبار

الفرضيّة الابتدائية $\mathbf{H}_o : \mu_1 = \mu_2$ عند مستوى الأهمّية $\alpha = 0.01$ ، وذلك لأنَّ مستوى الثقة المقدَّم $1-\alpha = 0.99$، ومن أجل استكمال دراسة هذا الاختبار وفقاً للحالات التي اتخذناها نجد ما يلي:

من أجل الطلب الأول: وبخصوص الفقرة:

أ- نلاحظ من مُعطيات العيّنات أنّه من المرجَّح أن يكون $\mu_1 < \mu_2$ ولذلك يفضَّل في مثل هذه الحالة اختيار الفرضيّة البديلة $\mathbf{H}_A : \mu_1 < \mu_2$ مع الأخذ بالحسبان أنَّ قيم الانحراف المعياري (ومن ثمَّ التباين) للمجتمعين معلومة حيث لدينا $\sigma_1 = \sigma_{1,o} = 25$ و $\sigma_2 = \sigma_{2,o} = 25$. عندئذ تُعطى منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ بوساطة المتباينة $z \leq -z_{1-\alpha}$، علماً أنَّ $z$ تُحسب بوساطة العلاقة [11,11] حيث لدينا:

$$z = \frac{(\bar{x} - \bar{y}) - a_o}{\sqrt{m\,\sigma_{1,o} + n\,\sigma_{2,o}}} \sqrt{n \cdot m} \quad = \frac{(480 - 505) - 0}{\sqrt{49 \times 25 + 36 \times 25}} \sqrt{49 \times 36} = -1$$

وبما أنَّ $z_{1-\alpha} = z_{0.99} \approx 2.33$ فإنَّه ينتج لدينا أنَّ:

$$-z_{1-\alpha} = -2.33 < -1 = z$$

ومن ثمَّ نجد أنَّ قيمة $z$ تقع في منطقة القبول للفرضيّة $\mathbf{H}_o$ ، ومن ثمّ يمكن اختيار أيّ من الشركتين لتوريد الدواء عند مستوى الأهمّية $\alpha = 0.01$ ، وأنَّ ما ظهر من فوارق في نتائج تحليل العيّنات هي فوارق ظاهرية لا أهمّية لها عند هذا مستوى الأهمّية المُعطى.

ب- سنأخذ الفرضيّة البديلة $\mathbf{H}_A : \mu_1 < \mu_2$ أيضاً، ولكن من أجل هذه الحالة قيم التباين للمجتمعين مجهولة $\sigma_1^2 = ?$ و $\sigma_2^2 = ?$. عندئذ تُعطى منطقة الرفض للفرضيّة $\mathbf{H}_o$ بوساطة المتباينة $t \leq t_{1-\alpha}^{[n+m-2]}$، علماً أنَّ $t$ تُحسب بوساطة العلاقة [11,12]، ولكن قبل استكمال دراسة هذا الاختبار علينا التحقُّق أولاً من تساوي التباين للمجتمعين وذلك من خلال تعيين 99% فترة ثقة للنسبة $\sigma_1^2 / \sigma_2^2$، علماً أنَّ هذه الفترة تُعطى بالفترة [10,34-b] حيث لدينا من جدول نوزيع فيشر ما يلي:

$$f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]} = f_{0.995}^{[48\,;\,35]} \approx \frac{2.63 + 2.42}{2} = 2.525$$

وبالتعويض في الفترة [10,34-b] كل بما يساويه يكون لدينا 99% فترة ثقة للنسبة $\dfrac{\sigma_1^2}{\sigma_2^2}$ هي:

$$\left( \frac{S_X^2}{S_Y^2} \cdot \left( f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]} \right)^{-1} , \; \frac{S_X^2}{S_Y^2} \cdot f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]} \right) = \left( \frac{400}{225}(2.525)^{-1} \, , \; \frac{400}{225}(2.525) \right) = (0.704 \, , \, 4.489)$$

وهكذا نجد أنَّ هذه الفترة تحوي العدد (1)، ومن ثمَّ لدينا $\sigma_1^2 = \sigma_2^2$ باحتمال قدره 0.99 ، ومن ثَمّ يمكننا متابعة دراسة الاختبار حيث لدينا من جدول توزيع ستودنت $t_{1-\alpha}^{[n+m-2]} = t_{0.99}^{[83]} = 2.372$، وبالتعويض كل بما يساويه في العلاقة [11,12] ينتج لدينا:

$$t = \frac{\left[(480 - 505) - 0\right] \sqrt{(36)(49)(36 + 49 - 2)}}{\sqrt{(49 + 36)\left[35 \times 400 + 48 \times 225\right]}} = -0.08$$

وهكذا نلاحظ أنَّ:

$$t_{1-\alpha}^{[n+m-2]} = -2.372 < -0.08 = t$$

ومن ثمَّ قيمة $t$ تقع في منطقة القبول للفرضيّة الابتدائية $\mathbf{H}_o$ ، وبالتالي يمكن اختيار أي من الشركتين لتوريد الدواء عند مستوى الأهمّية $\alpha = 0.01$ أيضاً، وأنَّ ما ظهر من فوارق في نتائج تحليل العيّنات في هذه الحالة هي فوارق ظاهرية أيضاً ولا أهمّية لها عند هذا المستوى من الأهمّية.

**من أجل الطلب الثاني:**

من أجل اختبار الفرضيّة القائلة بتساوي التباين لدى مجتمعي أدوية الشركتين نلاحظ أنّنا سنتوقع القبول لهذه الفرضيّة لأنّنا استنتجنا في الطلب السابق أنّ $\sigma_1^2 = \sigma_2^2$ باحتمال قدره 0.99، ولكن لتنفيذ الاختبار سنأخذ الفرضيّة الابتدائية $\mathbf{H}_o : \sigma_1^2 = \sigma_2^2$ مقابل الفرضيّة البديلة $\mathbf{H}_A : \sigma_1^2 \neq \sigma_2^2$ عند مستوى من الأهميّة $\alpha = 0.01$ . حيث نعلم أنّ منطقة الرفض للفرضيّة $\mathbf{H}_o$ تُعطى من خلال المتباينتين الآتيتين:

$$f \geq f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]} = f_{0.995}^{[48\,;\,35]} \approx 2.525 \qquad \& \qquad f \leq \left(f_{1-\frac{\alpha}{2}}^{[m-1\,;\,n-1]}\right)^{-1} = \left(f_{0.995}^{[48\,;\,35]}\right)^{-1} \approx \frac{1}{2.525} = 0.396$$

علماً أنّ $f$ تُحسب بوساطة العلاقة [11,13] حيث لدينا:

$$f = \frac{S_Y^2}{S_X^2} \cdot b_o = \frac{225}{400} \times 1 = 0.5625$$

فنلاحظ أنّ قيمة $f$ تقع في منطقة القبول للفرضيّة الابتدائية $\mathbf{H}_o$ (وهذا ما توقعناه مسبقاً)، ومن ثمّ نقبل بفرضية تساوي تباين أوزان المادة الفعّالة في كبسولات أدوية الشركتين وذلك عند مستوى الأهميّة $\alpha = 0.01$ .

## (١١,٦) اختبارات من أجل الفرق بين معلمتي مجتمعين برنوليين مستقلين

تقدم لنا المبرهنة الآتية (نقبلها دون برهان ) اختباراً لفرضيّة إحصائية متعلّقة بالفرق بين مَعْلمتي مجتمعين برنوليين مستقلّين.

**(١١,٦,١) مبرهنة**

ليكن $[\Omega_1\,,\,\mathbb{S}_1\,,\,\mathcal{P}_1]$ مجتمعاً موصوفاً من خلال متغيّر عشوائي $X$ خاضع للتوزيع $\mathbf{B}(p_1)$، ولنأخذ $\mathcal{X}_n$ عيّنة من هذا المجتمع. كذلك لنأخذ $[\Omega_2\,,\,\mathbb{S}_2\,,\,\mathcal{P}_2]$ مجتمعاً آخر مستقل عن $[\Omega_1\,,\,\mathbb{S}_1\,,\,\mathcal{P}_1]$ وموصوفاً من خلال متغيّر عشوائي $Y$ خاضع للتوزيع $\mathbf{B}(p_2)$، ولنأخذ $\mathcal{Y}_m$ عيّنة مسحوبة من هذا المجتمع الثاني، ولتكن $\mathbf{H}_o : p_1 - p_2 = b_o$ فرضيّة ابتدائية يُطلب اختبارها عند مستوى من الأهميّة $\alpha$، علماً أنّ $b_o$ ثابت حقيقي. إنّ منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ تُعيّن على النحو الآتي:

١- تُعطى منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ بوساطة المتباينة $z_{1-\frac{\alpha}{2}} \leq |z|$ إذا كانت الفرضيّة البديلة من النموذج:

$$\mathbf{H}_A : p_1 - p_2 \neq b_o \,.$$

٢- تُعطى منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ بوساطة المتباينة $z_{1-\alpha} \leq z$ إذا كانت الفرضيّة البديلة من النموذج:

$$\mathbf{H}_A : p_1 - p_2 > b_o$$

٣- تُعطى منطقة الرفض للفرضيّة الابتدائية $\mathbf{H}_o$ بوساطة المتباينة $-z_{1-\alpha} \geq z$ إذا كانت الفرضيّة البديلة من النموذج:

$$\mathbf{H}_A : p_1 - p_2 < b_o$$

علماً أنّ $z$ تُحسب بوساطة العلاقة الآتية:

$$Z = \frac{(\bar{p}_1 - \bar{p}_2) - b_o}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n} + \frac{1}{m}\right)}} \quad ; \bar{p} = \frac{k+\ell}{n+m} \qquad [11,14]$$

والعددين $k$ و $\ell$ هما عدد النجاحات في العينة $\mathcal{X}_n$ و $\mathcal{Y}_m$ على الترتيب.

**(١١,٦,٢) ملاحظة**

يُلاحظ أنّه في الحالة الخاصّة عندما يكون $b_o = 0$ فإنّ الاختبارات السابقة تؤول إلى اختبارات تتعلّق بتساوي مَعْلمتي مجتمعين برنوليين مستقلّين.

**(١١,٦,٣) مثال**

يقوم مُدرِّسان A وB بتصحيح دفاتر امتحانيه لمقرَّر الإحصاء وبشكل مستقل كل منهما عن الآخر. قمنا بأخذ عيّنة من الدفاتر المصحَّحة من قبل المدرس A وB بحجم $n = 225$ و $m = 200$ دفتراً امتحانياً على الترتيب، ولدى مراجعة التصحيح لهذه الدفاتر وجد أنَّ لدى A وB دفاتر تحوي أخطاءً في التصحيح عددها 20 و17 دفتراً على الترتيب. والسؤال هنا:

**هل تدلّ هذه النتائج أنَّ للمدرِّسين الكفاءة نفسها في دقّة التصحيح عند مستوى من الأهميّة قدره $\alpha = 0.01$ ؟**

الحـل: نلاحظ هنا أنَّ عمليات التصحيح للدفاتر الامتحانية هي تجارب برنولية (إمّا أن يكون الدفتر قد صُحِّح بدون أخطاء، أو أنَّ المدرس أخطأ في تصحيحه) مستقلّة (لأنَّ كل مدرِّس يصحِّح بشكل مستقل عن الآخر)، فإذا كان $\left[\Omega_1, \mathfrak{F}_1, \mathcal{P}_1 = \{P_{p_1}; 0 < p_1 < 1\}\right]$ هو المجتمع الممثِّل للدفاتر المصحَّحة من قبل المدرس A، و $\left[\Omega_2, \mathfrak{F}_2, \mathcal{P}_2 = \{P_{p_2}; 0 < p_2 < 1\}\right]$ هو المجتمع الممثِّل للدفاتر المصحَّحة من قبل المدرس B، فعندئذ للإجابة على السؤال المطروح علينا اختبار الفرضيّة الابتدائية $\mathbf{H}_o : p_1 = p_2$ مقابل الفرضيّة البديلة $\mathbf{H}_A : p_1 \neq p_2$ عند مستوى الأهميّة $\alpha = 0.01$.

الآن، من مُعطيات المسألة لدينا $\bar{p}_1 = \frac{20}{225} = 0.089$ وكذلك $\bar{p}_2 = \frac{17}{200} = 0.085$، ومن أجل $\alpha = 0.01$ لدينا:

$$z_{1-\frac{\alpha}{2}} = z_{0.995} \approx \frac{2.57 + 2.58}{2} = 2.575$$

وأمّا قيمة z فإنَّها تحسب من العلاقة [11,14] حيث لدينا:

$$z = \frac{(\bar{p}_1 - \bar{p}_2) - b_o}{\sqrt{\bar{p}(1-\bar{p})\left(\frac{1}{n} + \frac{1}{m}\right)}}$$

والتعويض عن كل رمز بما يساويه وتنفيذ العمليات الحسابية عليها ينتج لدينا الآتي:

$$z = \frac{(0.089 - 0.085) - 0}{\sqrt{0.087(1 - 0.087)\left(\frac{1}{225} + \frac{1}{200}\right)}} = 0.5182$$

ومن ثمَّ نجد أنَّ قيمة $z$ تقع في منطقة القبول للفرضيّة الابتدائية $\mathbf{H}_o$، وبالتالي نقبل أنَّ للمدرسين الدقّة نفسها في التصحيح وذلك عند مستوى الأهميّة $\alpha = 0.01$.

## (١١,٧) اختبارات إحصائية من أجل فرضيات غير معلمية

**Statistical Tests for Non-Parametric Hypothesis**

قدَّمنا فيما سبق بعض الاختبارات المَعْلميّة الأكثر شيوعاً واستخداماً، إلاّ أنّه توجد فرضيات إحصائية كثيرة تقع خارج إطار تلك الدراسة لعدم تعلّق الفرضيّة بمَعَالم المجتمع، كأن يكون الهدف من الاختبار الإحصائي هو التحقّق من أنَّ التوزيع الاحتمالي للمجتمع يوافق توزيعًا احتماليًّا معلومًا (اختبار فرضيات توفيق التوزيعات الاحتمالية). كذلك من الممكن أن يكون الهدف من الاختبار

الإحصائي التّحقُّق من استقلال توزيعين أو ظاهرتين لمجتمعين (اختبار فرضيات الاستقلال)، وقد يكون الهدف من ذلك التّحقُّق من التطابق في التوزيع لعدد منته من المجتمعات (اختبار فرضيات التجانس). وهناك الكثير من الاختبارات الإحصائية التي تتناول قطاعات أخرى في نظرية اختبارٌ الفرضّيات الإحصائية.

سنقوم فيما يلي بتقديم بعض الاختبارات غير المَعْلميّة مبتدئين بما يُعرف باسم **اختبارات جودة توفيق التوزيعات (التوزيعات الاحتمالية)** التي تنطوي تحت **الاختبارات غير المَعْلميّة**.

## (١ , ٧ , ١١) اختبارات جودة توفيق توزيع احتمالي Goodness of Distribution Fit Tests

لقد رأينا أنّه من أجل مجتمع $[\Omega , \mathbb{S}, \mathcal{P}]$ موصوف من خلال متغيِّر عشوائي $X$ دالّة كثافته الاحتمالية $f_X(\bullet;\theta)$ وبدالّة توزيـع احتمالية $F_X$ مجهولتين يُمكننا أن نشكِّل صورة تقريبية عن $F_X$ من خلال سـحب عيّنة $\mathfrak{X}_n$ من هذا المجتمع، ومن ثّم تعيين دالّة التوزيع العمليـّة للعيِّنة $F_{n,\omega}$ التي تتقارب من $F_X$ بشكل منتظم (بحسب مبرهنة غليفينكو)، كما أنَّ المضلَّع التكراري لهذه العيّنة يُعطينا عرضاً تقريبياً لدالّة الكثافة الاحتمالية $f_X(\bullet;\theta)$. إنَّ هذه التقريبات تدعونا إلى وضع فرضيّة حول نوع التوزيع الاحتمالي الذي يخضع له المجتمع، ومن ثّم اختبار هذه الفرضيّة عند مستوى معيِّن من الأهميّة $\alpha$ .

سنقدِّمها فيما يلي أحد اختبارات جودة التوفيق، ويُدعى **اختبار كاي مربَّع لجودة التوفيق**، وكان قد قدَّمه لأول مرَّة الإحصائي **بيرسون**.

### (١ , ١ , ٧ , ١١) اختبار كاي مربَّع لجودة التوفيق Chi-Squared Test for Goodness of Fit

إنَّ اختبار كاي مربَّع لجودة توفيق توزيع احتمالي ينطلق من الآتي:

بفرض أنّه لدينا $[\Omega , \mathbb{S}, \mathcal{P}]$ مجتمع موصوف من خلال متغيِّر عشوائي $X$ دالّة توزيعه الاحتمالية $F_X$ مجهولة، فعندئذ لاختبار إن كان توزيع هذا المجتمع يتوافق مع توزيع احتمالي $F_o$ (نعتقد أنّه مناسب لهذا المجتمع) عند مستوى من الأهميّة $\alpha$ نقوم بتنفيذ الخطوات الست الآتية:

١- نسحب عيّنة $\mathfrak{X}_n$ من هذا المجتمع.

٢- نضع الفرضيّة الابتدائية $\mathbf{H}_o : F_X \equiv F_o$ .

٣- تكون الفرضيّة البديلة في هذه الدراسة هي $\mathbf{H}_A : F_X \neq F_o$ .

٤- نقوم بتجزئة العيّنة $\mathfrak{X}_n$ في $m$ فئة منفصلة مثنى مثنى.

٥- نحسب قيمة الإحصاءة الآتي:

$$U = \sum_{k=1}^{m} \frac{\left(O_k - \mathbf{E}_k\right)^2}{\mathbf{E}_k} \qquad [11,15]$$

علماً أنَّ:

$O_k$ هو التكرار المُشَاهد للظاهرة لدى الفئة $k$ الذي يمثِّل التكرار المطلق للفئة $k$ ، حيث نلاحظ أنَّ $O_k$ متغيِّر عشوائي وقيمته تُعبِّر عن عدد المرَّات التي يأخذ فيها $X$ قيماً في الفئة $k$ .

$E_k$ هو التكرار المتوقَّع للفئة $k$ ويحسب بالعلاقة $\mathbf{E}_k = n \cdot p_k$ ، علماً أنَّ $p_k$ هو احتمال أن يأخذ المتغيِّر العشوائي قيمة في الفئة $k$ وذلك باستخدام التوزيع الاحتمالي $F_o$ ، وهكذا نلاحظ أنَّ $p_k$ هو التكرار النسبي المتوقع للفئة ذات الرقم $k$.

٦- تعيين عدد صحيح $r$ يُمثِّل عدد المَعَالم المُقدَّرة للتوزيع $F_o$ .

عندئذ بفرض أنَّ $u$ هي قيمة الإحصاءة $U$ فإنّنا:

أ - نقبل الفرضيّة الابتدائية $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha$ إذا كان $u < \chi_{1-\alpha}^{[k-r-1]}$ .

ب- نرفض الفرضيّة الابتدائية $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha$ إذا كان $u \geq \chi_{1-\alpha}^{[k-r-1]}$ .

**(١١,٧,١,١,١) ملاحظات**

١ - إنَّ التكرار المتوقَّع $\mathbf{E}_k$ يجب ألّا يكون صغيراً، إذ يُفضَّل أن يكون $\mathbf{E}_k \geq 5$ من أجل أي $k \ni \mathbb{N}_m$ ، وفي حال عدم تحقُّق المتباينة السابقة نقوم بدمج بعض الفئات بعضها مع البعض الآخر بحيث تَتحقَّق المتباينة السابقة.

٢ - إذا كان $X$ متغيِّراً عشوائياً متقطِّعاً، فعندئذ من الأفضل أن نقوم باختبار الفرضيّة الابتدائية:

$$\mathbf{H}_o : P_\theta(X = k) = P_{\theta_o}(X = k)$$

علماً أنَّ $\theta_o$ هي مَعْلَمة التوزيع التي قيد الاختبار، ويتم تقديرها من العيِّنة إن كانت مجهولة.

٣- من إيجابيات هذا الاختبار أنَّه يمكن تطبيقه على أي توزيع وحيد البُعد (بمتغيِّر واحد) حيث يمكن حساب دالَّة توزيعه التراكمية.

٤ - لهذا الاختبار سلبياتان هما:

أ- إنَّ قيمة الإحصاءة $U$ للاختبار تعتمد على جدولة البيانات، وهذه نقطة سلبية لكونها تتطلب الكثير من الجهد.

ب- يتطلب أن يكون حجم العيِّنة كبيرًا بقدر كافٍ لكي يكون التقريب مقبولاً (أو جيداً).

**(١١,٧,١,١,٢) مثال (تجربة جايجر ومارسيدن Geiger–Marsden experiment)**

لقد قام (بين عامي ١٩٠٨-١٩١٣) كل من الفيزيائيين الألمانيين جايجر Johannes Wilhelm Geiger (1882-1945) ومارسدن Sir Ernest Marsden (1889-1970) وتحت إدارة الفيزيائي النيوزيلاندي ريذيرفورد Ernest Rutherford (1871-1937) بتجربة حول عدد جسيمات ألفا الصادرة عن المادة المشعّة (Polonium) خلال 2608 فترات زمنية قصيرة طول كل منها 7.5 ثانية (انظر[24])، فوجدا النتائج الآتية على افتراض أنَّ $O_k$ هو المتغيِّر العشوائي الذي يرصد عدد الفترات الزمنية التي حصلا فيها على $k$ جسيم.

الجدول (١١,٢)

| $k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 or moor |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $O_k$ | 57 | 203 | 383 | 525 | 532 | 408 | 273 | 139 | 45 | 27 | 16 |

الآن ليكن $X$ متغيِّراً عشوائياً راصداً **لعدد** جسيمات ألفا الصادرة عن هذه المادة المشعّة خلال فترة زمنية قدرها 7.5 ثانية، ولنقم باختبار الفرضيّة القائلة إنَّ للمتغيِّر العشوائي $X$ توزيعًا بواسونيًا، وذلك عند مستوى من الأهميّة قدره $\alpha = 0.05$ .

الحـل: المطلوب اختبار الفرضيّة الابتدائية:

$$\mathbf{H}_o : P_\lambda(X = k) = P_{\lambda_o}(X = k) = \frac{\lambda_o^k}{k!} \mathrm{e}^{-\lambda_o} \quad ; k \in \mathbb{N}^o$$

عند مستوى من الأهميّة $\alpha = 0.05$ مقابل الفرضيّة البديلة:

$$\mathbf{H}_A : P_\lambda(X = k) \neq P_{\lambda_o}(X = k) = \frac{\lambda_o^k}{k!} \mathrm{e}^{-\lambda_o} \quad ; k \in \mathbb{N}^o$$

لدراسة هذا الاختبار علينا حساب قيمة المَعْلَمة $\lambda_o$ من العيّنة $\mathfrak{X}_n$ حيث نعلم من المثال /٣/ في (١٠-١-١-٥) أنَّ $\bar{X}$ هو مُقدِّر الأرجحيَّة العظمى لمَعْلَمة المجتمع البواسوني $\lambda$ (ويبرهن بسهولة أنّه غير منحاز أيضاً)، ومن ثمَّ نجد أنَّ:

$$\overline{x} = \frac{\sum_{k=0}^{10} k\,O_k}{\sum_{k=0}^{10} O_k} = \frac{10086}{2608} = 3.867 = \lambda_o$$

ومن ثَمّ تكون قيمة $p_0$ المقابلة لـ $k = 0$ هي:

$$p_0 = P_{\lambda_o}(X = 0) = \frac{(3.867)^0}{0!}\mathbf{e}^{-3.867} = 0.021$$

وكذلك تكون قيمة $p_1$ المقابلة لـ $k = 1$ هي:

$$p_0 = P_{\lambda_o}(X = 1) = \frac{(3.867)^1}{1!}\mathbf{e}^{-3.867} = 0.081$$

وبالمثل نجد بقية القيم لـ $p_k$ المقابلة لـ $k$ والمقدَّمة في الجدول الآتي: $p_k$

الجدول (١١,٣)

| $k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 or moor |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $p_k$ | 0.021 | 0.081 | 0.156 | 0.201 | 0.195 | 0.151 | 0.097 | 0.054 | 0.026 | 0.011 | 0.007 |

ومن ثَمّ يكون لـ $u$ (قيمة الإحصاءة $U$ المعطى بالعلاقة [11,15]) القيمة الآتية:

$$u = \sum_{k=0}^{10} \frac{\left(O_k - n\cdot p_k\right)^2}{n\cdot p_k} = 13.049$$

ومن جهة أخرى نعلم أنَّ لتوزيع بواسون مَعْلَمةً واحدةً تمَّ تقديره سابقاً، ولذلك يكون لدينا $r = 1$ ، وكذلك لدينا $m = 11$ (لأنَّه لدينا إحدى عشرة فئة منفصلة)، وأخيراً من جدول توزيع كاي مربَّع نجد $\chi_{1-\alpha}^{[k-r-1]} = \chi_{0.95}^{[9]} = 16.9$ ، ومن ثَمَّ نلاحظ أنَّ:

$$\chi_{0.95}^{[9]} = 16.9 > 13.049 = u$$

وهذا يعني أنَّه علينا القبول بالفرضية الابتدائية $\mathbf{H}_o$ القائلة إنَّ توزيع عدد الجسيمات الصادرة عن مادة البولونيوم Polonium خلال فترات زمنية طولها 7.5 من الثواني هو التوزيع البواسوني بمَعْلَمة $\lambda_o = 3.867$ ، وذلك عند مستوى الأهميّة $\alpha = 0.05$ .

### (١١,٧,١,٢) اختبار كلموغوراف – سميرنوف لجودة التوفيق Kolmogorov-Smirnov Test for Goodness of Fit

لقد قدّم كلٌّ من كلموغوراف والرياضياتي الروسي سميرنوف Nikolai Vasilyevich Smirnov (1900-1966) اختباراً لجودة توفيق توزيع احتمالي (وذلك في عام 1939) مستفيدين من انحرافات قيم دالَّة التوزيع العمليّة عن دالَّة التوزيع النظرية المفترض صحتها. حيث يبيِّن هذا الاختبار إن كان توزيع احتمالي $F_o$ ملائمًا لتوزيع هذا المجتمع أم لا، وذلك عند مستوى من الأهميّة قدره $\alpha$ .

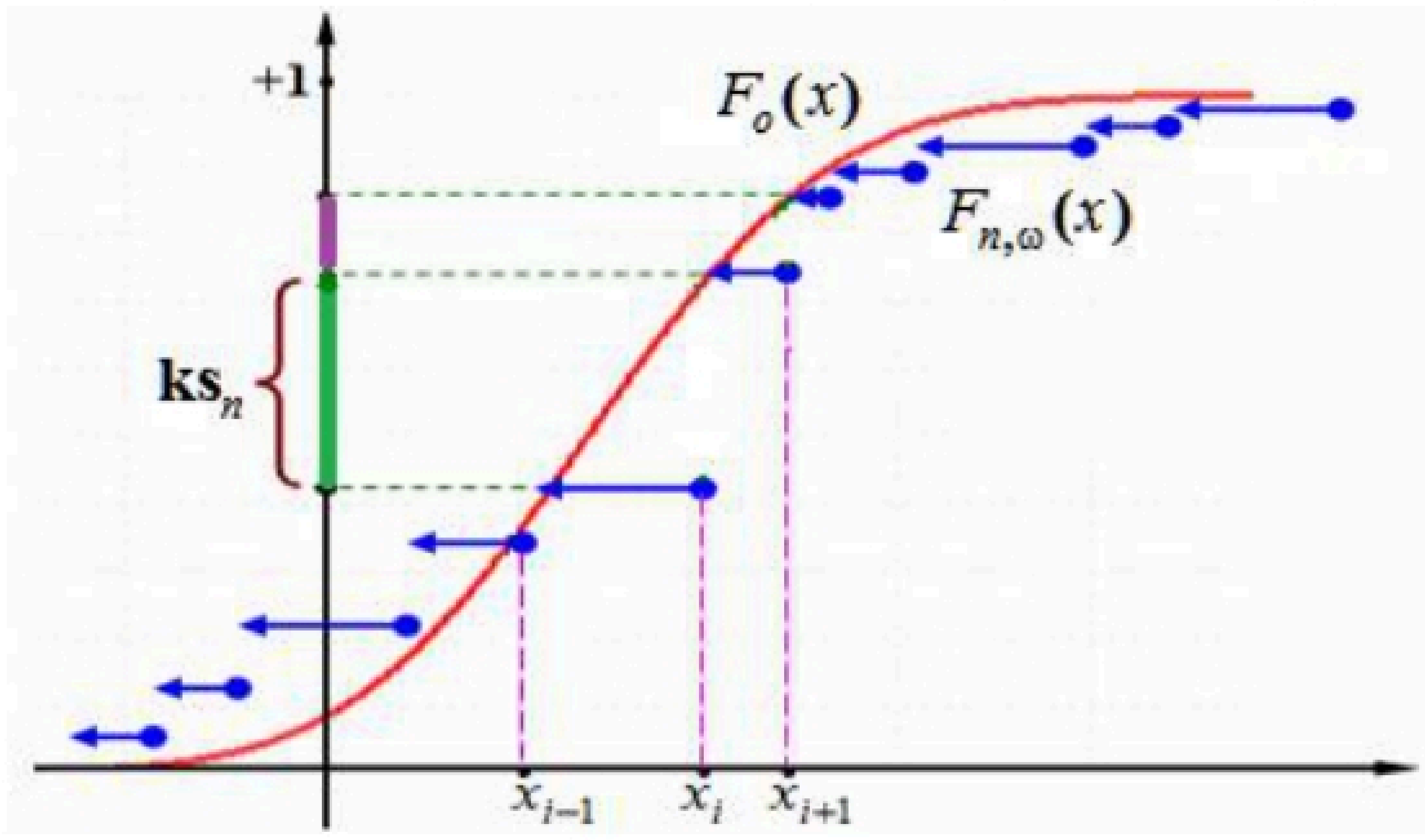
ولشرح هذا الاختبار سنفترض أنَّ $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمع موصوف من خلال متغيِّر عشوائي $X$ دالَّة توزيعه الاحتمالية $F_X$ مجهولة ولكنَّها مستمرَّة، وأنَّ $\mathcal{X}_n$ عيِّنة من هذا المجتمع، فعندئذ يقوم اختبار كلموغوراف– سميرنوف على تنفيذ البنود الآتية:

١- وضع الفرضيّة الابتدائية $\mathbf{H}_o : F_X \equiv F_o$ ، وفي حال رفض $\mathbf{H}_o$ عند مستوى من الأهميّة $\alpha$ علينا القبول بفرضية بديلة من الشكل $\mathbf{H}_A : F_X \neq F_o$ .

٢- من أجل حجم العيِّنة $n$ يُعرَّف إحصاءة $\mathbf{KS}_n$ تحسب قيمتها $\mathbf{ks}_n$ من خلال العلاقة الآتية:

$$\mathbf{ks}_n := \sup_{x\in\mathbb{R}} \left| F_{n,\omega}(x) - F_o(x) \right| \qquad [11,16\text{-a}]$$

علماً أنَّ $F_{n,\omega}$ هي دالَّة التوزيع العمليـَّة للعيِّنة $\mathfrak{X}_n$ وتحسب من العلاقة [9–8-a] أو [9–8-b].




الشـكل (١١,٨)

٣- تُعيَّن قيمة عددية سنرمز لها بـ $K_{1-\alpha}^{[n]}$ من **جدول** كلموغوراف- سميرنوف في آخر هذا الكتاب، وتُدعى **القيمة الحرجة** لاختبـار كلموغوراف– سميرنوف.

٤- يُقارن بين القيمتين $\mathbf{ks}_n$ و $K_{1-\alpha}^{[n]}$، فإذا كان:

أ- لدينا $K_{1-\alpha}^{[n]} < \mathbf{ks}_n$ فإنَّنا نرفض الفرضيّة الابتدائية $\mathbf{H}_o$ ونقبل بالفرضية البديلة $\mathbf{H}_A$ عند مستوى الأهمّية $\alpha$.

ب- لدينا $K_{1-\alpha}^{[n]} \geq \mathbf{ks}_n$ فإنَّنا نقبل الفرضيّة الابتدائية $\mathbf{H}_o$ عند مستوى الأهمّية $\alpha$.

### (١١,٧,١,٢,١) ملاحظات

١- بعض المراجع تستخدم العلاقة الآتية لحساب $\mathbf{ks}_n$:

$$\mathbf{ks}_n := \sqrt{n} \sup_{x \in \mathbb{R}} \left| F_{n,\omega}(x) - F_o(x) \right| \qquad [11,16\text{-b}]$$

٢- توجد في بعض المراجع جداول للقيم الحرجة تختلف عن التي سنقدِّمها في نهاية هذا الكتاب، وتلك الجداول خاصّة باستخدام اختبار كلموغوراف– سميرنوف عندما تكون دالَّة التوزيع النظرية هي دالَّة التوزيع الطبيعي فقط.

٣- لا يستخدم هذا الاختبار إلاَّ إذا كانت دالَّة التوزيع النظرية مستمرّة فقط.

٤- إنَّ هذا الاختبار يُعدّ أكثر حساسيّة (بالغ الدقّة) من أجل القيم القريبة من المتوسط من تلك القيم الأبعد التي تقع عند الأطراف (عند الذيلين)، وهذا يعني أنَّه يُعطي وزنًا أكبر للبيانات القريبة من المركز.

### (١١,٧,١,٢,٢) مثال

لدى دراسة تعيين التوزيع الزمني لانهيار عمل دارة إلكترونية لدى تعرضها لجهد كهربي أعلى من الجهد المسموح به أخضعت عيّنة بحجم 10 دارات من المنتج للدراسة، فوجدت المُعطيات الآتية مُقدَّرة بالثانية:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 0.5 | 1 | 3 | 3.6 | 0.3 | 4 | 0.1 | 0.2 | 5 | 1.2 |

ولنختبر الفرضيّة القائلة إنَّ التوزيع الزمني لانهيار الدارات هو التوزيع الأسّي، وذلك عند مستوى من الأهمّية قدره $\alpha = 0.10$.

الحـل: من أجل ذلك علينا اختبار الفرضيّة الابتدائية $\mathbf{H}_o : F_X(x) \equiv F_o(x) = \lambda_o \, \mathrm{e}^{-\lambda_o x} \, \mathbf{I}_{[0\,,\,\infty)}(x)$ مقابل الفرضيّة البديلة $\mathbf{H}_A : F_X(x) \neq F_o(x)$ عند مستوى الأهمّية $\alpha = 0.10$.

الآن، و لتنفيذ هذا الاختبار علينا تعيين قيمة المَعْلَمة $\lambda_o$ ، حيث يمكن إثبات أنَّ مُقدِّر مَعْلَمة المجتمع الأسّي هو $\hat{\lambda} = \frac{1}{\bar{X}}$ (استخدم طريقةالعزوم أو الأرجحيَّة العظمى للحصول على هذه النتيجة)، ومنه نجد أنَّ قيمة الإحصاءة $\hat{\lambda}$ تساوي $\hat{\lambda} = 0.529$ ، ومن ثمَّ سنأخذ $\lambda_o = 0.529$ من أجل $F_o(x)$ .

بعد ذلك لنقم ترتيب القيم التي حصلنا عليها من العيّنة تصاعدياً، ومن ثمَّ حساب القيم $F_{n,\omega}(x)$ و $F_o(x)$ فتكون النتائج كما في الجدول (١١ , ٤) الآتي، حيث نجد $\mathbf{ks}_n = 0.1955$ .

الخطوة التالية هي إيجاد قيمة $K_{1-\alpha}^{[n]}$ من جدول توزيع كلموغوراف، حيث نجد $K_{1-\alpha}^{[n]} = K_{0.90}^{[10]} = 0.369$ ، ومنه ينتج لدينا:

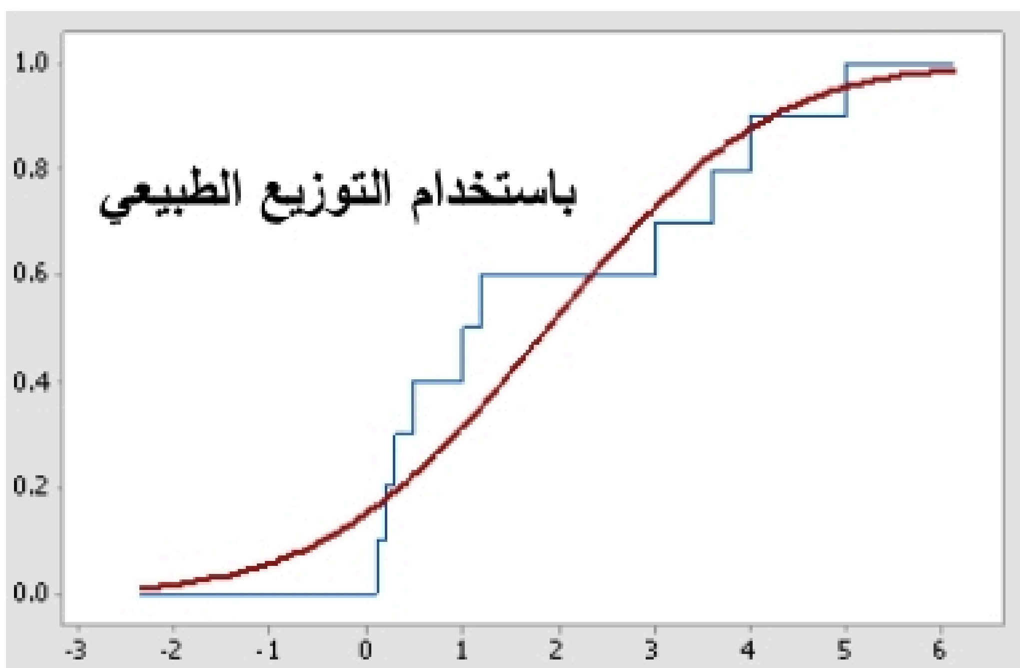
$$K_{1-\alpha}^{[n]} = 0.369 > 0.1955 = \mathbf{ks}_n$$

ومن ثمَّ ليس أمامنا إلاَّ القبول بالفرضية الابتدائية $\mathbf{H}_o$ عند مستوى الأهمّيّة $\alpha = 0.10$ .

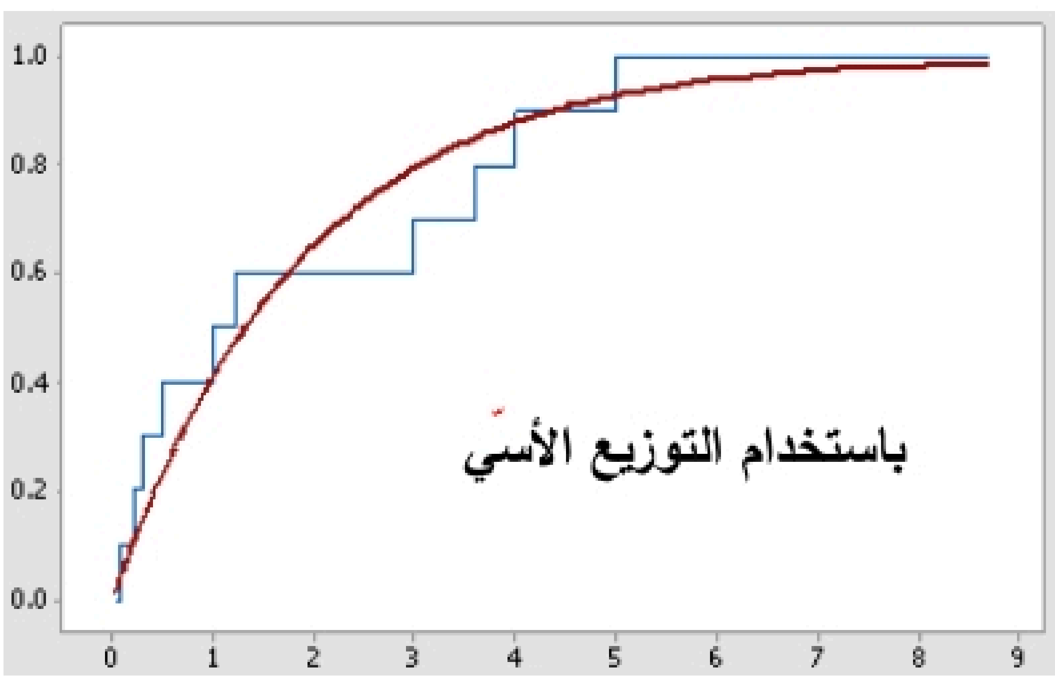
الجدول (١١ , ٤)

| الفئة | $x_i$ | $F_{n,\omega}(x_i)$ | $F_o(x_i)$ | $\lvert F_{n,\omega}(x_i) - F_o(x_i) \rvert$ |
|---|---|---|---|---|
| 1 | 0.1 | 0 | 0.052 | 0.0515 |
| 2 | 0.2 | 0.1 | 0.100 | 0.0004 |
| 3 | 0.3 | 0.2 | 0.147 | 0.0533 |
| 4 | 0.5 | 0.3 | 0.232 | 0.0676 |
| 5 | 1.0 | 0.4 | 0.411 | 0.0108 |
| 6 | 1.2 | 0.5 | 0.470 | 0.0300 |
| 7 | 3.0 | 0.6 | 0.795 | 0.1955 |
| 8 | 3.6 | 0.7 | 0.851 | 0.1511 |
| 9 | 4.0 | 0.8 | 0.879 | 0.0795 |
| 10 | 5.0 | 0.9 | 0.929 | 0.0290 |

والشكلان الآتيان يعرضان دالَّة التوزيع العمليَّة مع التوزيع الأسّي والطبيعي الملائم لها على الترتيب، وذلك باستخدام برنامج **Minitab**.




الشكل (١١ , ٩ ب)




الشكل (١١ , ٩ أ)

### (١١,٧,١,٣) اختبار أندرسون– دارلينغ لجودة التوفيق Anderson-Darling Test for Goodness of Fit

إنَّ اختبار أندرسون– دارلينغ الذي قدِّم عام 1952 من قبل الرياضياتي والإحصائي الأمريكي أندرسون Theodore Wilbur Anderson (1918-2016) والإحصائي الأمريكي دارلينغ Donald Allan Darling (1915-2014) يُعدّ بديلاً جيداً عن الاختبارات الإحصائية السابقة عندما نريد الكشف عن مدى ابتعاد توزيع العيّنة من التوزيع الطبيعي (أو من توزيع وايبول Weibull Distribution) كما هو الحال لدى اختبار كلموغوراف– سميرنوف، ولكنَّه يتميَّز عن هذا الاختبار الأخير أنَّه يُعطي وزناً أكبر لذيول التوزيع. كذلك هو اختبار إحصائي يبيِّن لنا ما إذا كان توزيع بيانات مُعطاة تتوافق مع توزيع احتمالي معيّن أيضاً، بمعنى أنَّه يمكن تطبيقه على جميع التوزيعات، ولكن إيجاد جدول القيم الحدّية له ليس بالأمر السهل، وهذا الجدول موجود من أجل التوزيع الطَّبيعي، والطَّبيعي اللوغاريتمي، والأسي، ووايبل، وتوزيع غامبل، والتوزيع المنطقي (اللوجستي).

بما أنَّ اختبار أندرسون– دارلينغ لجودة التوفيق (من أجل اختبار ملاءمة توزيع احتمالي ما لتوزيع هذا المجتمع) يستفيد من توزيع معيَّن في حساب القيم الحرجة فإنَّ ذلك ستمنحه ميزة إعطاء الاختبار حساسيّة أكثر، ولكن العيب في هذا الاختبار هو أنَّ القيم الحرجة يجب أن تحسب من أجل كل توزيع مستخدم على حدة.

فيما يلي سنقدِّم اختبار أندرسون– دارلينغ لتوفيق توزيع احتمالي ما عندما تكون دالَّة التوزيع النظرية $F_o$ هي دالَّة التوزيع الطبيعي $\mathbf{N}(\mu,\sigma)$، وفي حال عدم معرفة المَعْلمتين $\mu$ و $\sigma$ فإنَّه يتمّ تقديرهما من العيّنة $\mathfrak{X}_n$، فلو كان $[\Omega, \mathcal{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متغيِّر عشوائي $X$ دالَّة توزيعه $F_X$ مجهولة، و $\mathfrak{X}_n$ عيّنة من هذا المجتمع، فعندئذ يقوم اختبار أندرسون– دارلينغ على تنفيذ البنود الآتية:

١- وضع الفرضيّة الابتدائية $\mathbf{H}_o : F_X \equiv F_o$، وفي حال رفض $\mathbf{H}_o$ عند مستوى من الأهميّة $\alpha$ علينا القبول بالفرضية البديلة: $\mathbf{H}_A : F_X \neq F_o$، علماً أنَّ $F_o$ هي دالَّة التوزيع النظرية التي نفترض أن تكون صحيحة.

٢- نقوم بترتيب قيم البيانات الناتجة عن العيّنة تصاعدياً ليصبح لها العرض $x_1 < x_2 < \ldots < x_i < \ldots < x_n$.

٣- تُعرَّف إحصاءة $D$ بالعينة $\mathfrak{X}_n$ قيمتها (سنرمز لها بـ **ad**) تحسب باستخدام العلاقة الآتية:

$$\mathbf{ad} := -n + \frac{1}{n}\sum_{i=1}^{n}(1-2i)\left[\mathbf{ln}F_o(x_i) + \mathbf{ln}\left[1 - F_o(x_{n-i+1})\right]\right] \quad [11,17]$$

٤- تُحسب القيمة الحرجة لهذا الاختبار بالعلاقة الآتية:

$$\mathbf{c} = 0.752\left(1 + \frac{0.75}{n} + \frac{2.25}{n^2}\right)^{-1} \quad [11,18]$$

٥- تُقارن القيمة **ad** مع **c**، فإذا كان:

أ- لدينا $\mathbf{c} < \mathbf{ad}$ فعندئذ ترفض الفرضيّة الابتدائية $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha$.

ب- لدينا $\mathbf{c} \geq \mathbf{ad}$ فعندئذ تقبل الفرضيّة الابتدائية $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha$.

#### (١١,٧,١,٣,١) ملاحظة

بما أنَّ تعيين قيم $F_o(x_i)$ يعتمد على التوزيع الطبيعي ذي المَعْلمتين $\mu_o$ و $\sigma_o$ فإنَّ تنفيذ الاختبار يدوياً سيتطلب جهداً كبيراً ومطولاً، ولذلك يفضَّل استخدام البرامج الإحصائية لحساب **ad** ومن ثمّ مقارنتها مع القيمة الحرجة **c**.

**مثال (١١,٧,١,٣,٢)**

بالعودة إلى المثال السابق لننظر إن كان التوزيع الزمني لانهيار عمل دارة إلكترونية لدى تعرضها لجهد كهربي أعلى من الجهد المسموح به خاضع للتوزيع الطبيعي أم لا، وذلك عند مستوى من الأهمّية قدره $\alpha = 0.05$ ؟

**الحل**: لدينا من مُعطيات المثال البيانات الآتية مُقدَّرة بالثانية:

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 1.2 | 5 | 0.2 | 0.1 | 4 | 0.3 | 3.6 | 3 | 1 | 0.5 |

ولنختبر الفرضيّة الابتدائية:

$$\mathbf{H}_o : F_X(x) \equiv F_o(x) = \frac{1}{\sigma_o\sqrt{2\pi}}\ \mathbf{e}^{-\frac{1}{2}\left(\frac{x-\mu_o}{\sigma_o}\right)^2}$$

مقابل الفرضيّة البديلة $\mathbf{H}_A : F_X(x) \neq F_o(x)$ وذلك عند مستوى الأهمّية قدره $\alpha = 0.05$ .

لتنفيذ هذا الاختبار علينا تعيين قيم المَعْلمتين $\mu_o$ و $\sigma_o$ ، حيث نعلم أنَّ مُقدَّرا مَعْلمتي المجتمع الطبيعي $\mu$ و $\sigma$ هما $\hat{\mu} = \bar{X}$ و $\hat{\sigma} = S$ (مقدَّرات غير منحازة أيضاً)، ومن المعطيات التي لدينا نجد أنَّ $\hat{\mu} = \bar{X} = 1.89$ وكذلك $\hat{\sigma} = S = 1.83$ .

الآن من أجل حساب **ad** سنستخدم الجدول الآتي لتسهيل العمليات الحسابية:

| $i$ | $x_i$ | $a_i$ | $b_i$ | $c_i$ | $d_i$ | $e_i$ | $f_i$ | $g_i$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1.2 | 0.098 | -2.319 | 0.149 | 0.851 | -0.161 | -2.48 | 2.480 |
| 2 | 4.5 | 0.723 | -0.324 | 0.108 | 0.892 | -0.114 | -0.437 | 1.312 |
| 3 | 0.2 | 0.195 | -1.636 | 0.126 | 0.874 | -0.135 | -1.771 | 8.854 |
| 4 | 0.1 | 0.215 | -1.535 | 0.220 | 0.780 | -0.248 | -1.784 | 12.48 |
| 5 | 4.0 | 0.352 | -1.045 | 0.186 | 0.814 | -0.205 | -1.250 | 11.25 |
| 6 | 0.3 | 0.177 | -1.731 | 0.352 | 0.648 | -0.433 | -2.165 | 23.81 |
| 7 | 3.6 | 0.220 | -1.514 | 0.215 | 0.785 | -0.243 | -1.757 | 22.84 |
| 8 | 3.0 | 0.130 | -2.038 | 0.195 | 0.805 | -0.217 | -2.255 | 33.82 |
| 9 | 1.0 | 0.108 | -2.230 | 0.723 | 0.277 | -1.286 | -3.515 | 59.76 |
| 10 | 0.5 | 0.149 | -1.904 | 0.098 | 0.902 | -0.104 | -2.007 | 38.14 |
| Sum | 18.4 | | | | | | | 214.8 |
| Average | 1.84 | | | | | | | 21.48 |
| S.D. | 1.738 | | | | | | | |

علماً أنّنا استخدمنا الرموز الآتية للمحافظة على حدود الجدول ومظهره:

$$a_i \triangleq F_o(x_i) \quad \& \quad b_i \triangleq \ln F_o(x_i) \quad \& \quad c_i \triangleq F_o(x_{n-i+1}) \quad \& \quad d_i \triangleq 1 - F_o(x_{n-i+1})$$

$$e_i \triangleq \ln\left[1 - F_o(x_{n-i+1})\right] \quad \& \quad f_i := b_i + e_i \quad \& \quad g_i := (1-2i)\, f_i$$

فيكون لدينا ما يلي:

$$\mathbf{ad} = -n + \frac{1}{n}\sum_{i=1}^{n}(1-2i)\left[\ln F_o(x_i) + \ln\left[1 - F_o(x_{n-i+1})\right]\right] = 11.48$$

وأمّا القيمة الحرجة **c** فنجدها تساوي:

$$\mathbf{c} = 0.752\left(1 + \frac{0.75}{n} + \frac{2.25}{n^2}\right)^{-1} = 0.752\left(1 + \frac{0.75}{10} + \frac{2.25}{100}\right)^{-1} = 0.685$$

ومن ثَمّ يكون لدينا $\mathbf{ad} = 11.48 > 0.685 = \mathbf{c}$ ، ولذلك علينا القبول بالفرضية الابتدائية $\mathbf{H}_o$ عند مستوى الأهمّية $\alpha = 0.05$ .

لقد قمنا فيما سبق بتقديم اختبارات لفرضيات تتعلّق بمسألة توافق توزيع احتمالي مع توزيع مجتمع إحصائي مُعطى. فيما يلي سنتابع تقديم اختبارين آخرين ولكن ليس من أجل مجتمعين إحصائيين فقط وإنّما من أجل أي عدد منته من المجتمعات.

### (١١,٧,٢) اختبارات إحصائيّة من أجل تجانس التوزيعات Homogeneity Tests of Distributions

لنفترض أنّه لدينا $[\Omega_1, \mathbb{S}_1, \mathcal{P}_1]$، $[\Omega_2, \mathbb{S}_2, \mathcal{P}_2]$،...و $[\Omega_k, \mathbb{S}_k, \mathcal{P}_k]$ مجتمعات مستقلّة موصوفة من خلال متغيّرات عشوائية $X_1$، $X_2$،...و $X_k$ على الترتيب، ولنفترض أنّنا نرغب باختبار الفرضيّة الابتدائية الآتية عند مستوى من الأهميّة $\alpha$ :

$$\mathbf{H}_o : F_{X_1}(x) \equiv F_{X_2}(x) \equiv \dots \equiv F_{X_k}(x)$$

مقابل الفرضيّة البديلة:

$\mathbf{H}_A$ : التوزيعات $F_{X_1}$، $F_{X_2}$،...و $F_{X_k}$ ليست جميعها متطابقة

فعندئذ يُقال عن هذا الاختبار إنَّه **اختبار التجانس للمجتمعات** $[\Omega_1, \mathbb{S}_1, \mathcal{P}_1]$، $[\Omega_2, \mathbb{S}_2, \mathcal{P}_2]$،...و $[\Omega_k, \mathbb{S}_k, \mathcal{P}_k]$.

في هذا الكتاب سندرس حالتين فقط، وهما اختبار التجانس للمجتمعات الطبيعية ذات التباين المتساوي، وكذلك اختبار التجانس للمجتمعات البرنولية.

#### (١١,٧,٢,١) اختبار من أجل تساوي المتوسطات لمجتمعات طبيعية مستقلّة

لتكن $[\Omega_1, \mathbb{S}_1, \mathcal{P}_1]$، $[\Omega_2, \mathbb{S}_2, \mathcal{P}_2]$،...و $[\Omega_k, \mathbb{S}_k, \mathcal{P}_k]$ مجتمعات طبيعية مستقلّة تبايناتها متساوية، ولنفترض أنّنا نرغب باختبار الفرضيّة الابتدائية $\mathbf{H}_o : \mu_1 = \mu_2 = \dots = \mu_k$ عند مستوى من الأهميّة قدره $\alpha$ مقابل الفرضيّة البديلة الآتية:

$$\mathbf{H}_A : \exists ! \, i, j \in \mathbb{N}_k \text{ with } i \neq j \text{ and } \mu_i \neq \mu_j$$

والتي تكتب على النحو الآتي أيضاً:

$\mathbf{H}_A$ : $\mu_1$، $\mu_2$،...و $\mu_k$ ليست جميعها متساوية

فعندئذ للوصول إلى القرار الخاص بهذا الاختبار نقوم بما يلي:

١- نسحب عيّنات عشوائية من المجتمعات $[\Omega_1, \mathbb{S}_1, \mathcal{P}_1]$، $[\Omega_2, \mathbb{S}_2, \mathcal{P}_2]$،...و $[\Omega_k, \mathbb{S}_k, \mathcal{P}_k]$ بحجم $n_1$، $n_2$،...و $n_k$ على الترتيب، ثمّ نخضع هذه العيّنات للدراسة المطلوبة، ولنفرض أنَّ كل عيّنة يمكن أن تخضع إلى $m$ نوع من الدراسات (أو الظواهر، أو المُعَالِجات)، وهذا يعني أنَّ كل عيّنة يمكن تجزئتها إلى $m$ فئة، وكذلك نتائج هذه الدراسة قدّمت وفقاً للجدول الآتي:

الجدول (١١,٥)

| | | الدراسة (أو الظاهرة، أو المُعَالِجة) | | | | المجموع |
|---|---|---|---|---|---|---|
| | | 1 | 2 | ...... | $m$ | |
| العيّنة | 1 | $n_{11}$ | $n_{12}$ | ...... | $n_{1m}$ | $n_{1\bullet} = n_1$ |
| | 2 | $n_{21}$ | $n_{22}$ | ...... | $n_{2m}$ | $n_{2\bullet} = n_2$ |
| | $\vdots$ | $\vdots$ | $\vdots$ | $\vdots \; \vdots$ | $\vdots$ | $\vdots$ |
| | $k$ | $n_{k1}$ | $n_{k2}$ | ...... | $n_{km}$ | $n_{k\bullet} = n_k$ |
| Total | | $n_{\bullet 1}$ | $n_{\bullet 2}$ | ...... | $n_{\bullet m}$ | $n$ |

علماً أنَّ $n_{ij}$ هي قيمة المتغيّر العشوائي $N_{ij}$ الذي يرصد عدد العناصر من العيّنة $n_i$ التي تُحقّق الدراسة $j$، وكذلك لدينا:

$$n_{j\bullet} = \sum_{\ell=1}^{m} n_{j\ell} \;\; ; j \in \mathbb{N}_k \qquad \& \qquad n_{\bullet\ell} = \sum_{j=1}^{k} n_{j\ell} \;\; ; \ell \in \mathbb{N}_m \qquad \& \qquad n = \sum_{j=1}^{k} n_{j\bullet} = \sum_{\ell=1}^{m} n_{\bullet\ell}$$

٢- نقوم بحساب قيمة إحصاءة $F$ مُعطى بالعلاقة الآتية:

$$F = \frac{k\,(k-1)}{m\,(m-1)} \cdot \frac{\sum_{j=1}^{m}\left(m\sum_{i=1}^{k} N_{ij} - n\right)^2}{\sum_{i=1}^{k}\sum_{j=1}^{m}\left(k\,N_{ij} - n_{\bullet j}\right)^2} \qquad [11,19]$$

٣- نعيِّن القيمة الحرجة للاختبار $f_{1-\alpha}^{[m-1;m(k-1)]}$ من جدول توزيع فيشر $\mathbf{F}(m;n)$ .

٤- نقارن قيمة الإحصاءة $F$ مع القيمة الحرجة $f_{1-\alpha}^{[m-1;m(k-1)]}$، فإذا كان:

أ- لدينا $F \geq f_{1-\alpha}^{[m-1;m(k-1)]}$ فإنّنا نرفض الفرضيّة $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha$ ، ونقرّ بقبول الفرضيّة البديلة $\mathbf{H}_A$ .

ب- لدينا $F < f_{1-\alpha}^{[m-1;m(k-1)]}$ فإنّه ليس أمامنا إلاّ القبول بالفرضية $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha$ .

إنَّ هذا النوع من الاختبارات يستخدم بكثرة في مجال تصميم وتحليل التجارب Design and Analysis of Experiments.

**مثال (١١،٧،٢،١،١)**

في مدينة يوجد أربعة مخابز مركزية $B_1$ ، $B_2$ ، $B_3$ و $B_4$ لإنتاج نوعين من الخبز (الصامولي والرغيف)، ويعمل كل منها بشكل مستقل عن الآخر. فإذا افترضنا أنَّ عملية الإنتاج لهذه الخطوط تتبع التوزيع الطبيعي بمتوسط $\mu_1$ ، $\mu_2$ ، $\mu_3$ و $\mu_4$ وبانحراف معياري $\sigma_1$ ، $\sigma_2$ ، $\sigma_3$ و $\sigma_4$ على الترتيب علماً أنَّ $\sigma = \sigma_4 = \sigma_3 = \sigma_2 = \sigma_1$ ، وقمنا بسحب عيّنات عشوائية من إنتاج هذه المخابز (مقدَّرة بالكيلو غرام) بحجم $n_1 = 121$ و $n_2 = 144$ و $n_3 = 169$ و $n_4 = 256$ على الترتيب، موزعة على كل نوع كما في الجدول الآتي:

الجدول (١١،٦)

| المخبز | خبز رغيف | خبز صامولي | المجموع |
|---|---|---|---|
| $B_1$ | 70 | 51 | 121 |
| $B_2$ | 100 | 44 | 144 |
| $B_3$ | 120 | 49 | 169 |
| $B_4$ | 200 | 56 | 256 |
| sum | 490 | 200 | 690 |

ولنبيّن إن كان متوسط الإنتاج لهذه المخابز متساو عند مستوى من الأهميّة $\alpha = 0.05$ أم لا.

الحـل: من أجل ذلك علينا اختيار الفرضيّة الابتدائية $\mathbf{H}_o : \mu_1 = \mu_2 = \mu_3 = \mu_4$ مقابل الفرضيّة البديلة الآتية:

$\mathbf{H}_A :$ المتوسطات $\mu_1$ ، $\mu_2$ ، $\mu_3$ و $\mu_4$ ليست جميعها متساوية

وذلك عند مستوى من الأهميّة $\alpha = 0.05$ .

إنَّ اختبار هذه الفرضيّة يتطلب منّا حساب قيمة الإحصاءة $F$ المعطاة بالعلاقة [11,19]، علماً أنّه لدينا $m = 2$ و $k = 4$ ، فنجد ما يلي:

$$F = \frac{k\,(k-1)}{m\,(m-1)} \cdot \frac{\sum_{j=1}^{m}\left(m\sum_{i=1}^{k} n_{ij} - n\right)^2}{\sum_{i=1}^{k}\sum_{j=1}^{m}\left(k \cdot n_{ij} - n_{\bullet j}\right)^2} = \frac{4(4-1)}{2(2-1)} \cdot \frac{\sum_{j=1}^{m}\left(2\sum_{i=1}^{k} n_{ij} - 690\right)^2}{\sum_{i=1}^{k}\sum_{j=1}^{m}\left(4 \cdot n_{ij} - n_{\bullet j}\right)^2}$$

ومن معطيات المثال نجد:

$$\sum_{j=1}^{m}\left(2\sum_{i=1}^{k}n_{ij}-690\right)^2=(980-690)^2+(400-690)^2=168200$$

$$\sum_{i=1}^{k}\sum_{j=1}^{m}\left(4\cdot n_{ij}-n_{\bullet j}\right)^2=(4\times 70-490)^2+(4\times 51-200)^2+\dots+$$

$$+(4\times 200-490)^2+(4\times 56-200)^2=149584$$

ومن ثَمّ ينتج لدينا:

$$F=\frac{12}{2}\cdot\frac{168200}{149584}=\frac{2018400}{299168}=6.747$$

ومن جدول توزيع فيشر نجد $f_{1-\alpha}^{[m-1;m(k-1)]}=f_{0.95}^{[1;6]}=5.99$، ومن ثَمّ يكون لدينا:

$$f_{1-\alpha}^{[m-1;m(k-1)]}=5.99<6.747=F$$

وهذا يعني أنّنا نقرّ برفض الفرضيّة الابتدائية القائلة بتساوي متوسطات الإنتاج لهذه المخابز ونقبل بالفرضية البديلة $H_A$ عند مستوى الأهمّية $\alpha=0.05$ .

### (١١,٧,٢,٢) اختبار من أجل تساوي مَعَالم مجتمعات برنولية مستقلّة

لتكن $[\Omega_1,\mathbb{S}_1,\mathcal{P}_1]$، $[\Omega_2,\mathbb{S}_2,\mathcal{P}_2]$،...و $[\Omega_k,\mathbb{S}_k,\mathcal{P}_k]$ مجتمعات برنولية بمَعَالم $p_1$، $p_2$،...و $p_k$ على الترتيب، ولنفترض أنّنا نرغب في اختبار الفرضيّة الابتدائية $\mathbf{H}_o: p_1=p_2=\dots=p_k$ عند مستوى من الأهمّية قدره $\alpha$ مقابل الفرضيّة البديلة الآتية:

$\mathbf{H}_A$ : المَعَالم $p_1$، $p_2$،...و $p_k$ ليست جميعها متساوية **or** $\mathbf{H}_A: \exists! \, i,j\in\mathbb{N}_k \text{ with } i\neq j \text{ and } p_i\neq p_j$

فعندئذ للوصول إلى القرار الخاص بهذا الاختبار نقوم بما يلي:

١ – نقوم بسحب عيّنات عشوائية من المجتمعات $[\Omega_1,\mathbb{S}_1,\mathcal{P}_1]$، $[\Omega_2,\mathbb{S}_2,\mathcal{P}_2]$،...و $[\Omega_k,\mathbb{S}_k,\mathcal{P}_k]$ بحجم $n_1$، $n_2$، ...و $n_k$ على الترتيب، ثمّ نخضع هذه العيّنات للدراسة المطلوبة، فعندئذ من أجل كل $i\in\mathbb{N}_k$ ستتوزع نتائج كل عيّنة $n_i$ في فئتين إحداهما عناصرها تُحقّق النجاح، وليكن حجمها $n_{i1}$، والأخرى عناصرها تفشل في تحقيق المطلوب سنفترض حجمها يساوي $n_{i2}$، ومن ثمّ يُمكننا أن نصبّ نتائج هذه الدراسة في الجدول الآتي:

الجدول (١١,٧)

| | لعيّنة | | | | المجموع |
|---|---|---|---|---|---|
| | 1 | 2 | ...... | $k$ | |
| عدد العناصر التي حقّقت النجاح | $n_{11}$ | $n_{21}$ | ...... | $n_{k1}$ | $n_{\bullet 1}$ |
| عدد العناصر التي لم تحقّق النجاح | $n_{12}$ | $n_{22}$ | ...... | $n_{k2}$ | $n_{\bullet 2}$ |
| **Total** | $n_1$ | $n_2$ | ...... | $n_k$ | $n$ |

علماً أنّ $n_{ij}$ هي قيمة المتغيّر العشوائي $N_{ij}$ الذي يرصد عدد عناصر العيّنة ذات الرقم $i$ والمحقّقة للخاصة ذات الرقم $j$ مع $j\in\{1,2\}$ و $i\in\mathbb{N}_k$، وكذلك لدينا:

$$n_{\bullet 1}=\sum_{i=1}^{k}n_{i1} \qquad \& \qquad n_{\bullet 2}=\sum_{i=1}^{k}n_{i2} \qquad \& \qquad n=n_{\bullet 1}+n_{\bullet 2}=\sum_{i=1}^{k}n_i$$

٢- نقوم بحساب قيمة إحصاءة $\chi$ مُعطى بالعلاقة الآتية (وسنرمز لقيمتها بـ $\chi$ أيضاً):

$$\chi = n \cdot \sum_{i=1}^{k} \sum_{j=1}^{2} \left[ \frac{\left(n \cdot n_{ij} - n_{\bullet j}\, n_i\right)^2}{n_{\bullet j}\, n_i} \right] \qquad [11,20]$$

٣- نستخرج $\chi_{1-\alpha}^{[k-1]}$ التي تمثّل القيمة الحرجة لهذا الاختبار من جدول توزيع كاي مربّع.

٤- نقارن قيمة الإحصاءة $\chi$ مع القيمة الحرجة $\chi_{1-\alpha}^{[k-1]}$، فإذا كان:

أ- لدينا $\chi_{1-\alpha}^{[k-1]} \le \chi$ فعندئذ نرفض الفرضيّة $\mathbf{H}_o$، ومن ثمّ القبول بالفرضية البديلة $\mathbf{H}_A$ عند مستوى الأهميّة $\alpha$.

ب- لدينا $\chi_{1-\alpha}^{[k-1]} > \chi$ فعندئذ ليس أمامنا إلاّ القبول بالفرضية الابتدائية $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha$.

**مثال (١١,٧,٢,٢,١)**

ذُكر أنَّ مرضًا جرثوميًّا يصيب الأغنام يجتاح منطقة رعوية تعيش فيها قطعان من الأغنام، وكإجراء وقائي ضد هذا المرض الجرثومي تَمَّ إعطاء اللقاح إلى خمسة قطعان من الأغنام $H_1$، $H_2$، $H_3$، $H_4$ و $H_5$ (قطيع الأغنام الواحد لا يقل عن **500** رأس) منفصلة ومتباعدة بعضها عن بعض (أي لدينا مجتمعات مستقلّة)، وبعد فترة من الزمن سحبت عيّنات عشوائية من هذه القطعان الخمسة بحجم $n_1 = 85$، $n_2 = 77$، $n_3 = 95$، $n_4 = 118$ و $n_5 = 95$ على الترتيب، فوجد أنَّ عدد الأغنام المُصابة في هذه العيّنات على الترتيب هي 3، 2، 2، 13 و7.

**السؤال هنا:** هل تشير هذه النتائج إلى أنَّ نسب الإصابات في هذه القطعان الخمسة متساوية وذلك عند مستوى من الأهميّة قدره $\alpha = 0.05$، وهل يبقى القرار نفسه من أجل $\alpha = 0.01$ ؟

الحـل: للإجابة على هذا السؤال علينا اختيار الفرضيّة الابتدائية $\mathbf{H}_o : p_1 = p_2 = p_3 = p_4 = p_5$ مقابل الفرضيّة البديلة:

$\mathbf{H}_A$ : المَعالم $p_1$، $p_2$، ...،و $p_k$ ليست جميعها متساوية

وذلك عند مستوى من الأهميّة $\alpha = 0.05$، وبعد ذلك سنعيد دراسة الاختبار نفسه من أجل مستوى الأهميّة $\alpha = 0.01$.

الآن، ومن أجل تبسيط تنفيذ هذا الاختبار سنصب نتائج دراستنا في الجدول الآتي:

الجدول (١١,٨)

| | | القطيع | | | | | المجموع |
|---|---|---|---|---|---|---|---|
| | | $H_1$ | $H_2$ | $H_3$ | $H_4$ | $H_5$ | |
| الحالة | سليمة | 82 | 75 | 93 | 105 | 88 | 443 |
| | مصابة | 3 | 2 | 2 | 13 | 7 | 27 |
| Total | | 85 | 77 | 95 | 118 | 95 | 470 |

ومن ثمّ يكون للإحصاءة $\chi$ المعطى بالعلاقة [11,20] القيمة الآتية:

$$\chi = n \cdot \sum_{i=1}^{5} \sum_{j=1}^{2} \left[ \frac{\left(n \cdot n_{ij} - n_{\bullet j}\, n_i\right)^2}{n_{\bullet j}\, n_i} \right] = 11.025$$

حيث تُعيَّن القيمة الحرجة $\chi_{1-\alpha}^{[k-1]}$ من جدول توزيع كاي مربّع، فنجدها من أجل $\alpha = 0.05$ تساوي:

$$\chi_{1-\alpha}^{[k-1]} = \chi_{0.95}^{[4]} = 9.488$$

ومن ثمَّ ينتج لدينا أنَّ:

$$\chi_{1-\alpha}^{[k-1]} = 9.488 < 11.025 = \chi$$

وهذا يعني رفض الفرضيّة الابتدائية $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha = 0.05$، أي إنَّ نسب الإصابات في هذه القطعان الخمسة ليست متساوية عند مستوى من الأهميّة $\alpha = 0.05$.

أمّا لتعيين القيمة الحرجة $\chi_{1-\alpha}^{[k-1]}$ من أجل $\alpha = 0.01$ فنجدها من جدول توزيع كاي مربّع أنّها تساوي:

$$\chi_{1-\alpha}^{[k-1]} = \chi_{0.99}^{[4]} = 13.277$$

ومن ثمَّ ينتج لدينا:

$$\chi_{1-\alpha}^{[k-1]} = 13.277 > 11.025 = \chi$$

وفي هذه الحالة ليس أمامنا إلاّ أن نقبل بالفرضية الابتدائية $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha = 0.01$، أي إنَّ نسب الإصابات في هذه القطعان الخمسة متساوية عند مستوى من الأهميّة $\alpha = 0.01$.

### (١١,٧,٣) اختبار استقلال ظاهرتين (أو متغيِّرين) Independence Test of two Phenomenon (or Variables)

ليكن $[\Omega, \mathbb{S}, \mathcal{P}]$ مجتمعاً موصوفاً من خلال متَّجه عشوائي ثنائي البُعد $\mathbb{X}_2 = (X_1, X_2)$، فعندئذ كل عنصر من عناصر المجتمع سيكون خاضعاً لصفتين (أو توزيعين) أحدهما ذو صلة بتوزيع $X_1$ والآخر ذو صلة بتوزيع $X_2$، فلو افترضنا أنَّ أي عنصر من المجتمع يمكن له أن يتواجد في إحدى الحالات $S_1$، $S_2$، ...و $S_k$ وفقاً للظاهرة الممثَّلة بـ $X_1$، وكذلك يمكن له أن يوجد في إحدى الحالات $Z_1$، $Z_2$، ...و $Z_\ell$ وفقاً للظاهرة الممثَّلة بـ $X_2$. عندئذ يمكننا صبَّ نتائج دراسة عيّنة من هذا المجتمع في الجدول الآتي:

الجدول (١١,٩)

| | | حالات المتغيِّر $X_1$: $S_1$ | $S_2$ ......... | $S_k$ | المجموع |
|---|---|---|---|---|---|
| حالات المتغيِّر $X_2$ | $Z_1$ | $n_{11}$ | $n_{12}$ ......... | $n_{1k}$ | $n_{1\bullet}$ |
| | $Z_2$ | $n_{21}$ | $n_{22}$ ......... | $n_{2k}$ | $n_{2\bullet}$ |
| | $\vdots$ | $\vdots$ | $\vdots$ $\vdots$ | $\vdots$ | $\vdots$ |
| | $Z_\ell$ | $n_{\ell 1}$ | $n_{\ell 2}$ ......... | $n_{\ell k}$ | $n_{\ell\bullet}$ |
| **Total** | | $n_{\bullet 1}$ | $n_{\bullet 2}$ ......... | $n_{\bullet k}$ | $n$ |

وهذا الجدول يُدعى **جدول الاقتران**، علماً أنَّ $n_{ij}$ هي قيمة المتغيِّر العشوائي $N_{ij}$ الذي يرصد عدد العناصر من العيّنة التي تُحقِّق الحالة $S_i$ و $Z_j$ معاً، وكذلك لدينا:

$$n_{i\bullet} = \sum_{j=1}^{k} n_{ij} \quad ; i \in \mathbb{N}_\ell \qquad \& \qquad n_{\bullet j} = \sum_{i=1}^{\ell} n_{ij} \quad ; j \in \mathbb{N}_k$$

والآن من أجل اختبار الفرضيّة الابتدائية:

$\mathbf{H}_o$ : $X_2$ مستقل عن $X_1$

مقابل الفرضيّة البديلة:

$\mathbf{H}_A$ : $X_2$ غير مستقل عن $X_1$

وذلك عند مستوى من الأهمّية $\alpha$، فإنّه علينا تنفيذ الخطوات الآتية:

١- نقوم بحساب قيمة الإحصاءة الآتية (وسنرمز لقيمتها بـ $u$):

$$U = \sum_{j=1}^{k} \sum_{j=1}^{\ell} \frac{\left(N_{ij} - \mathbf{E}_{ij}\right)^2}{\mathbf{E}_{ij}} \qquad [11,21\text{-a}]$$

علماً أنَّ المقادير $\mathbf{E}_{ij}$ تُدعى **التكرارات المتوقَّعة** Expected Frequencies وتحسب بالعلاقة الآتية:

$$\mathbf{E}_{ij} = \frac{n_{i\bullet} \cdot n_{\bullet j}}{n} \quad ; i \in \mathbb{N}_\ell \ , j \in \mathbb{N}_k \qquad [11,21\text{-b}]$$

٢- نعيِّن القيمة $\chi^{[v]}_{1-\alpha}$ من جدول توزيع كاي مربّع، علماً أنَّ $v = (k-1)\cdot(\ell-1)$.

٣- نقارن بين القيمتين $u$ و $\chi^{[v]}_{1-\alpha}$، فإذا كان:

أ- لدينا $\chi^{[v]}_{1-\alpha} \le u$، فعندئذ تُرفض الفرضيّة الابتدائية $\mathbf{H}_o$ ونقبل بالفرضية البديلة $\mathbf{H}_A$ عند مستوى الأهمّية $\alpha$.

ب- لدينا $\chi^{[v]}_{1-\alpha} > u$ فإنّه علينا القبول بالفرضية الابتدائية $\mathbf{H}_o$ عند مستوى الأهمّية $\alpha$.

### مثال (١١,٧,٣,١)

أُخذت عيّنة بحجم 160 جهازاً من أجهزة الحواسيب المستعملة بمعدل 15 ساعة عمل يومياً في درجات حرارة متباينة، فوجدنا أن الأجهزة التي اختل عملها مُعطاة كما في الجدول الآتي:

الجدول (١٠,١١. أ)

| | | درجة الحرارة المئوية التي عمل فيها الحاسوب | | | المجموع |
|---|---|---|---|---|---|
| | | 25 | 35 | 40 | |
| مدّة استخدام الجهاز بالشهر | 6 | 3 | 5 | 7 | 15 |
| | 12 | 5 | 9 | 15 | 29 |
| | 24 | 8 | 13 | 20 | 41 |
| | 36 | 14 | 23 | 38 | 75 |
| Total | | 30 | 50 | 80 | 160 |

فهل تدل هذه النتائج على أنَّ تأثير درجة الحرارة على الأجهزة غير مستقلّة عن عمر الجهاز عند مستوى الأهمّية قدره $\alpha = 0.05$، وهل يبقى القرار نفسه من أجل $\alpha = 0.01$ ؟

الحـل: من أجل الإجابة على هذا السؤال علينا اختبار الفرضيّة الابتدائية:

$\mathbf{H}_o$ : **تأثير تغيُّرات درجة الحرارة على الجهاز غير مستقلّة عن عمر الجهاز**

مقابل الفرضيّة البديلة:

$\mathbf{H}_A$ : **تأثير تغيُّرات درجة الحرارة على الجهاز لا علاقة لها بعمر الجهاز**

وذلك عند مستوى الأهمّية $\alpha = 0.05$.

الآن، من أجل حساب قيمة الإحصاءة $U$ المُعطاة بالعلاقة [11,21-a] سنقوم بحساب التكرارات المتوقَّعة $\mathbf{E}_{ij}$ باستخدام العلاقة [11,21-b] فنجدها كما في الجدول الآتي الذي يُدعى **جدول التكرارات المتوقعة** Table of Expected Frequencies:

الجدول (١٠ ,١١ .ب)

| | | قيم $i$ | | | المجموع |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | |
| قيم $j$ | 1 | 2.813 | 4.688 | 7.500 | قيم $\mathbf{E}_{ij}$ |
| | 2 | 5.438 | 9.063 | 14.500 | |
| | 3 | 7.688 | 12.813 | 20.500 | |
| | 4 | 14.063 | 23.438 | 37.500 | |

ولحساب القيم $\dfrac{\left(n_{ij}-\mathbf{E}_{ij}\right)^2}{\mathbf{E}_{ij}}$ سنستخدم الجدول الآتي:

الجدول (١٠ ,١١ .ج)

| | | قيم $i$ | | | المجموع |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | |
| قيم $j$ | 1 | 0.01243 | 0.02077 | 0.03333 | 0.06653 |
| | 2 | 0.03528 | 0.00044 | 0.01724 | 0.05296 |
| | 3 | 0.01266 | 0.00273 | 0.0122 | 0.02759 |
| | 4 | 0.00028 | 0.00819 | 0.00667 | 0.01513 |
| | | قيم $(n_{ij}-\mathbf{E}_{ij})^2 / \mathbf{E}_{ij}$ | | | 0.16193 |

فيكون لدينا ما يلي:

$$u=\sum_{j=1}^{4}\sum_{i=1}^{3}\frac{\left(n_{ij}-\mathbf{E}_{ij}\right)^2}{\mathbf{E}_{ij}}=0.162$$

علماً أنَّ عدد درجات الحرية $v$ يساوي $v=(k-1)(\ell-1)=2\times 3=6$، وأمّا لتعيين القيمة الحرجة $\chi^{[v]}_{1-\alpha}$ فنجدها باستخدام جدول توزيع كاي مربّع أنَّها:

أ- من أجل $\alpha=0.05$ تساوي $\chi^{[v]}_{1-\alpha}=\chi^{[6]}_{0.95}=12.592$، ومن ثمَّ ينتج لدينا:

$$\chi^{[v]}_{1-\alpha}=12.592>u=0.162$$

وفي هذه الحالة ليس أمامنا إلاّ القبول بالفرضية الابتدائية $\mathbf{H}_o$ عند مستوى الأهمّية $\alpha=0.05$، أي إنَّ تأثير تغيُّرات درجة الحرارة على الجهاز غير مستقلّة عن عمر الجهاز عند مستوى من الأهمّية $\alpha=0.05$.

ب- من أجل $\alpha=0.01$ تساوي $\chi^{[v]}_{1-\alpha}=\chi^{[6]}_{0.99}=16.812$، ومن ثمَّ ينتج لدينا أنَّ:

$$\chi^{[v]}_{1-\alpha}=16.812>u=0.162$$

وفي هذه الحالة علينا القبول بالفرضية الابتدائية $\mathbf{H}_o$ أيضاً عند مستوى الأهمّية $\alpha=0.01$.

فيما يلي نقدِّم بعض التطبيقات على استخدام برنامج Minitab في مجال اختبار الفرضيّات.

## (١١,٨) تطبيقات

**Applications**

### (١١,٨,١) تطبيق

قامت شركة مقاولات بنصب أعمدة كهرباء على طريق عام، وقد زعمت أنَّها على ثقة قدرها 95% أنَّ متوسط انزياحات أعمدتها التي قامت بنصبها معدوم (يساوي الصفر)، بانحراف معياري لا يتجاوز ميليمترًا واحدًا فقط. للتحقُّق من زعم هذه الشركة اختيرت عيِّنة عشوائية بحجم 49 عموداً من الأعمدة التي نصبتها هذه الشركة، ومن ثمَّ قيست مقادير الانزياح (الأخطاء) لهذه الأعمدة عن الموضع المخصَّص لها باستخدام أجهزة دقيقة، فكانت لدينا البيانات الآتية مقدَّرة بالميليمتر (وسنطلق عليها اسم Error في هذه الدراسة):

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2.1 | -0.5 | -1.6 | -0.8 | -0.7 | 2.1 | -1.1 | 1.2 | -1.4 | 0.4 | 0.7 | -2.4 | |
| -0.9 | -0.7 | -1.6 | -0.7 | -1.5 | -0.2 | -0.8 | -0.6 | -0.1 | -0.6 | -0.8 | -0.2 | |
| 1.8 | 0.2 | 0.7 | 1.9 | 0.3 | 0.9 | 0.7 | -0.9 | 1.0 | 0.2 | -0.2 | 0.8 | |
| -0.4 | 0.1 | 0.4 | -2.4 | 2.4 | 0.3 | -0.2 | -0.2 | -0.2 | -0.1 | 0 | 0.7 | 0.4 |

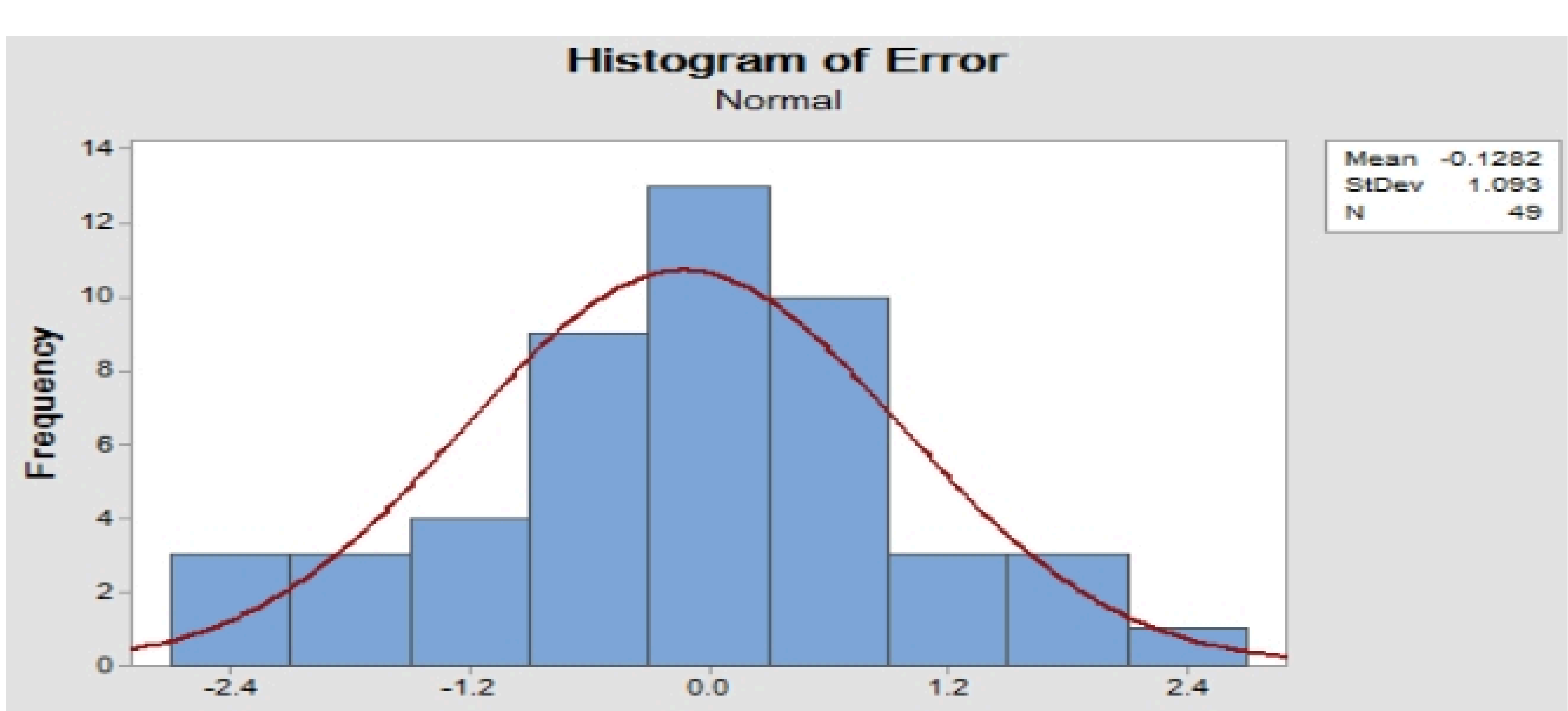
علماً أنَّ إشارة الناقص تشير إلى انزياح العمود نحو اليسار في حين تشير إشارة الموجب إلى انزياح العمود نحو اليمين.

والآن، لنختبر الفرضيّة الابتدائية القائلة إنَّ متوسط مجتمع الأخطاء المرتكبة في نصب الأعمدة يساوي الصفر عند مستوى الأهمّية $\alpha = 0.05$.

إذاً علينا اختبار الفرضيّة الابتدائية $\mathbf{H}_o : \mu = \mu_o = 0$ مقابل الفرضيّة البديلة $\mathbf{H}_A : \mu \neq \mu_o = 0$ وذلك عند مستوى من الأهمّية قدره $\alpha = 0.05$.

من أجل ذلك لننظر أولاً إن كان توزيع هذه البيانات طبيعيًّا أو قريبًا من الطبيعي، فنجد باستخدام برنامج Minitab أنَّ شكل مدرَّجه التكراري المقدَّم في الشكل الآتي قريب من الطبيعي (مع مقارنته مع التوزيع الطبيعي)، ويمكن الوصول إليه وفقاً للخطوات الآتية (شرحت في نهاية الفصل السابق):

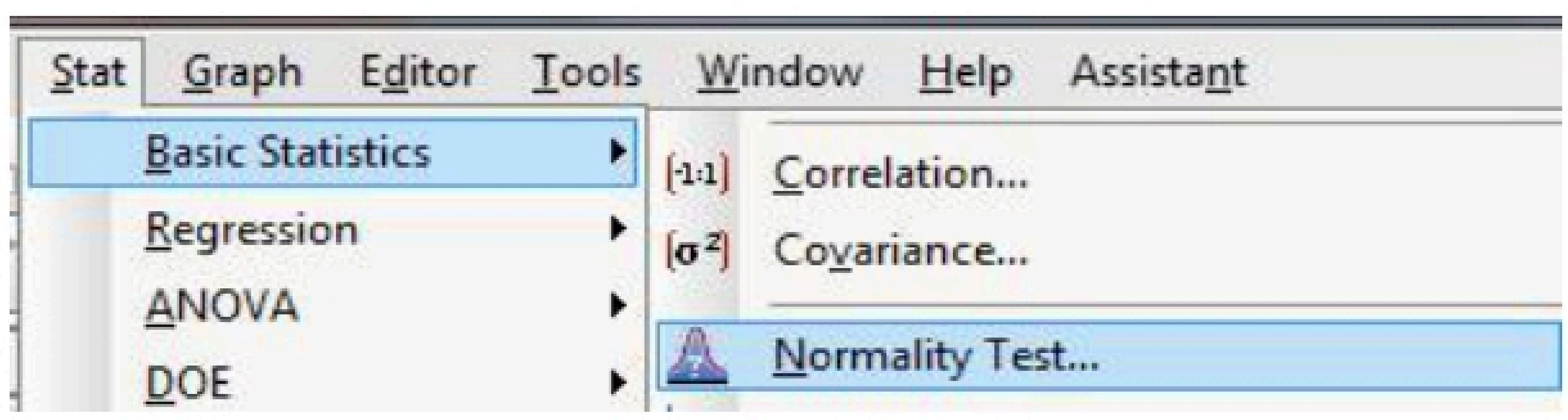
**Graph ⟶ Histogram ⟶ With Fit ⟶ Graph variables ⟶ In free space for Graph Variables Insert Error**
**ومن ثمَّ OK** ، يمكن الدخول إلى الأيقونات الأخرى قبل الضغط على موافق (OK) لإجراء بعض التعديلات حسب طبيعة الدراسة ومتطلباتها.



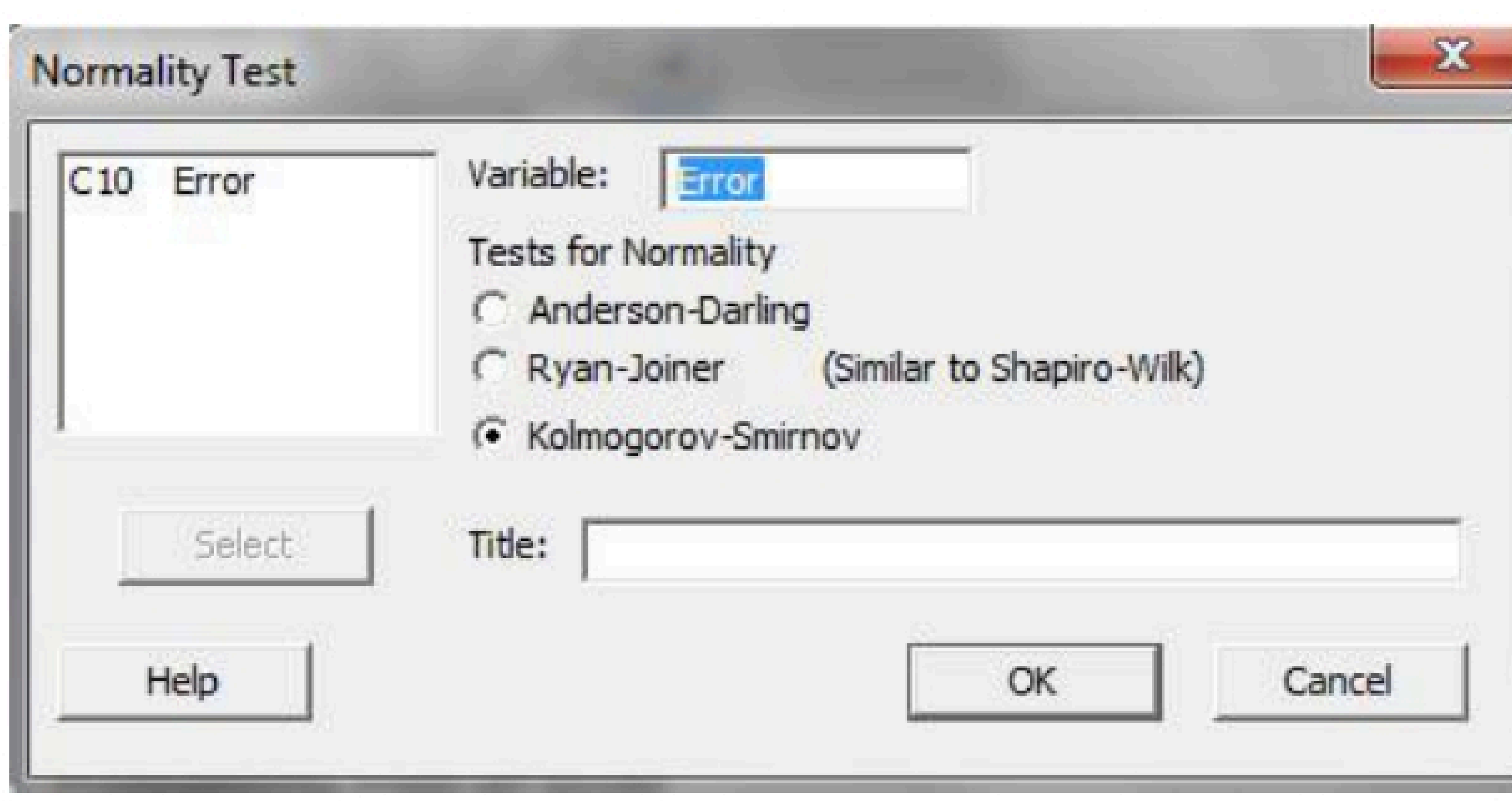

حيث نلاحظ بالفعل أنَّ المدرَّج التكراري يبدي تقارباً جيداً مع منحنى دالَّة الكثافة الاحتمالية للتوزيع الطبيعي، ويزودنا هذا العرض

بقيمة المتوسط والانحراف المعياري للبيانات المقدّمة، حيث لدينا $\overline{x}=-0.1282\,mm$ و $s=1.093\,mm$، وأمّا من أجل اختبار إن كانت هذه البيانات تقبل التوزيع الطبيعي كتوزيع مناسب فعلاً فإنّنا نجد باستخدام اختبار كلموغوراف– سميرنوف في برنامج Minitab ما يلي:

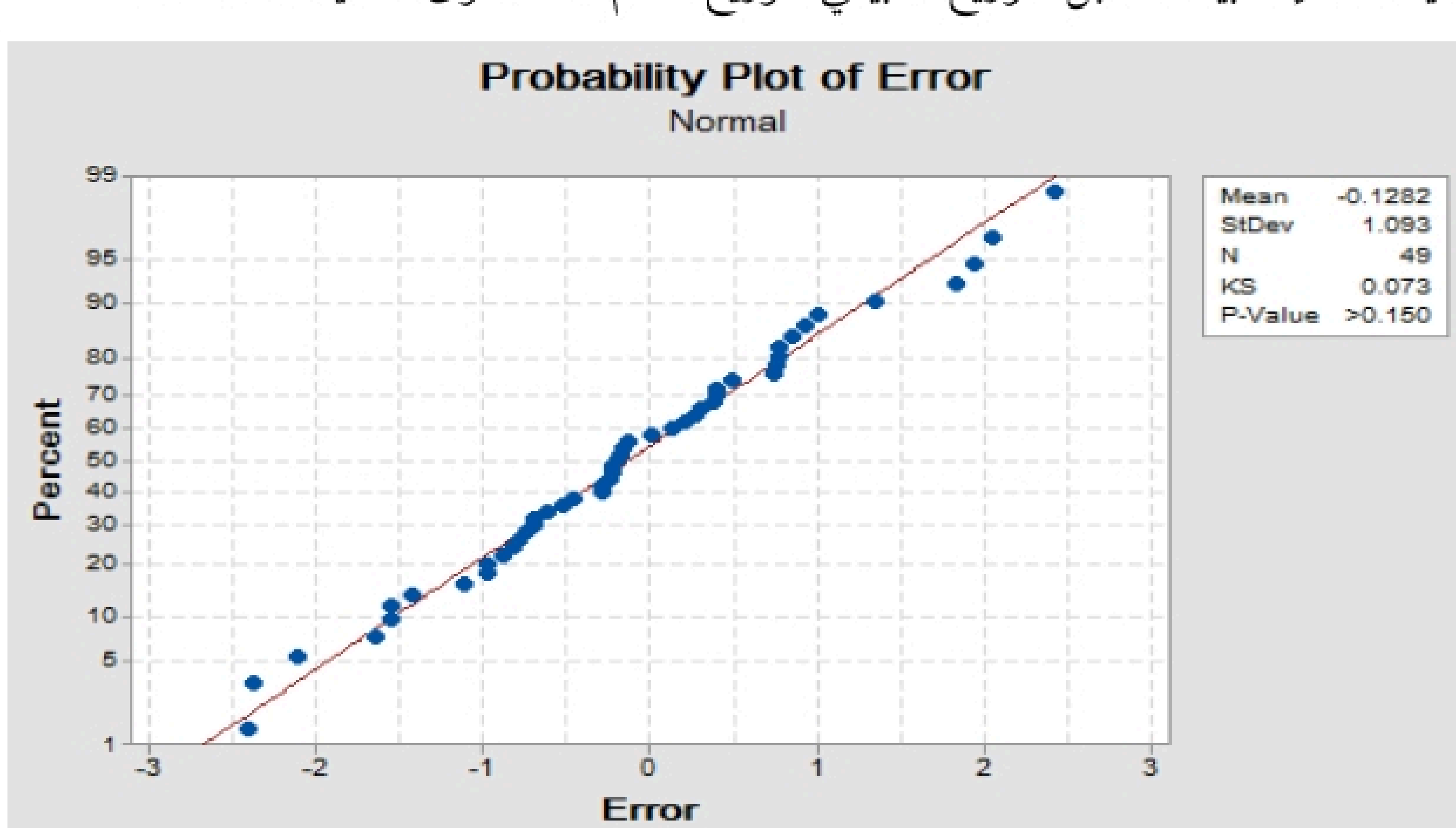
**Stat → Basic Statistics → Normality Test → Insert Error in free space for Variable → For Hypothesis test activation Kolmogorov-Smirnov → OK.**




ثمّ الخطوة التالية:




فنحصل على الشكل الآتي الذي يوضّح أنّ البيانات تقبل التوزيع الطبيعي (المعياري من أجل مثالنا هذا) عند مستوى الأهمّية $\alpha=0.05$، وذلك لأنّ القيمة المحسوبة **P-value** أكبر من 0.150 وهذه القيمة الأخيرة أكبر من مستوى الأهمّية $\alpha=0.05$ التي تعني قبول الفرضيّة القائلة إنّ البيانات تقبل التوزيع الطبيعي كتوزيع ملائم عند مستوى الأهمّية $\alpha=0.05$ .

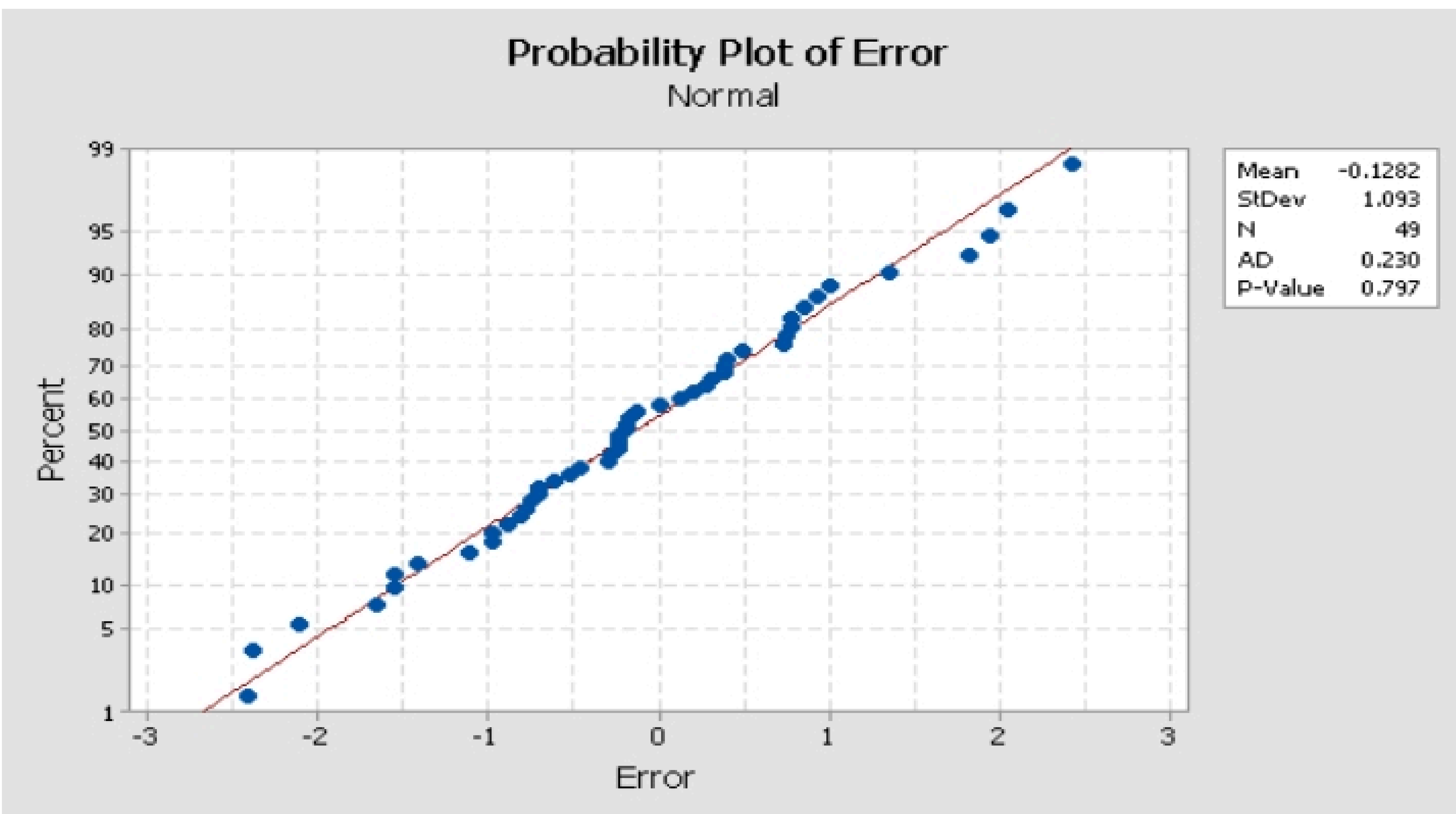


علماً أنَّ **P- value** (ويدعوها البعض "القيمة المحسوبة P " في حين يدعوها آخرون "القيمة الاحتمالية P") تُحدِّد مدى ملاءمة رفض الفرضيّة الابتدائية في دراسة اختبار الفرضيّات، وأمّا مجموعة القيم التي تنتمي إليها هذه القيمة فهي الفترة $[0, 1]$، ومعظم البرامج الإحصائية تقدِّمها في تقاريرها لدى اختبار الفرضيّات الإحصائية. أمّا تفسير هذه القيمة فيتمّ على النحو الآتي:

كلّما صَغُرَت القيمة الاحتمالية P أدى ذلك إلى صغَر الخطأ في احتمال رفض الفرضيّة الابتدائية، وعلى العموم فإذا كانت قيمة **P- value** أصغر من قيمة مستوى الأهمّية $\alpha$ فإنَّ ذلكَ يشير إلى رفض الفرضيّة الابتدائية $\mathbf{H}_o$ عند مستوى الأهمّية $\alpha$، وفي حال العكس يكون القبول لـ $\mathbf{H}_o$ .

الآن، و بالعودة إلى مثالنا يمكننا مُلاحظة تقارب النقاط الممثِّلة لقيم العيّنة من المستقيم الناتج عن خطية التوزيع الطبيعي المرسوم على لوحة مقسَّمة تُدعى لوحة التوزيع الطبيعي، حيث منحنى دالَّة التوزيع الطبيعي يبدو مستقيماً على هذه اللوحة (بسبب عملية الخطيّة على العلاقة التي تعطي دالَّة الكثافة للتوزيع والتي وضّحنا إياها لدى دراسة الانحدار الخطي)، وهذا يعني أنَّ البيانات تتوزَّع طبيعياً على وجه التقريب عند مستوى الأهمّية المفترض.

كذلك لو استخدمنا اختبار أندرسون– دارلينغ فإنَّنا سنحصل على النتيجة نفسها حيث لدينا الشكل الآتي لهذا الاختبار .



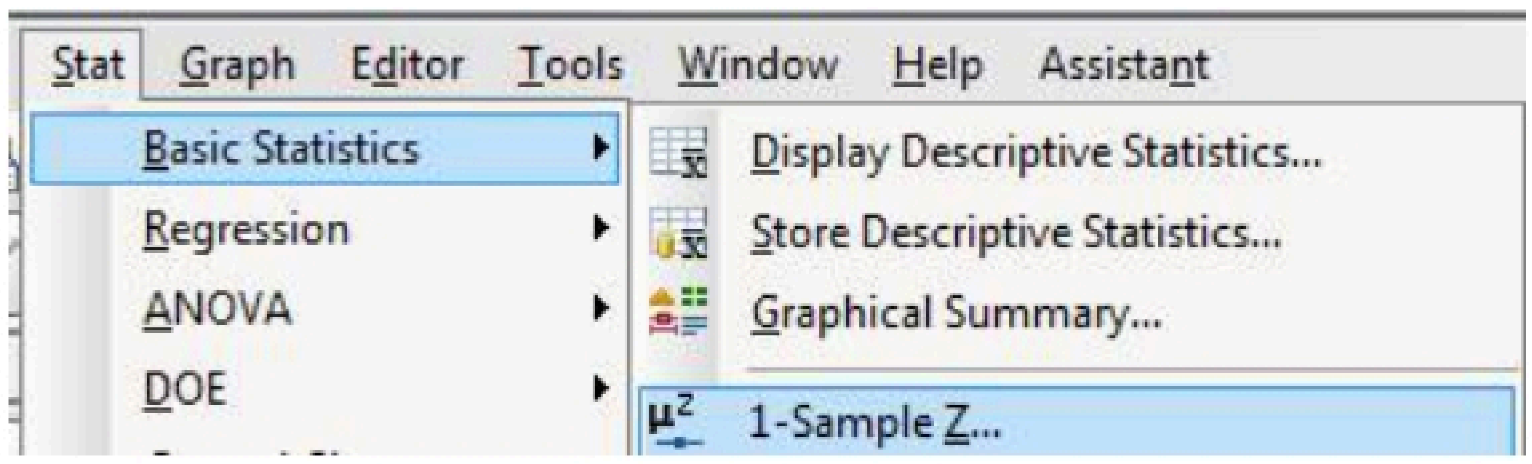

حيث نلاحظ هنا أنَّ $\alpha = 0.05 < 0.797 = $ **P– value** والذي يشير إلى قبول الفرضيّة الابتدائية القائلة إنَّ للبيانات الُمعطاة توزيعًا طبيعيًّا عند مستوى الأهمّية $\alpha = 0.05$ .

والآن من أجل اختبار الفرضيّة الابتدائية المقدَّمة سابقاً نعلم أنَّه علينا أن نميِّز بين حالتين هما:

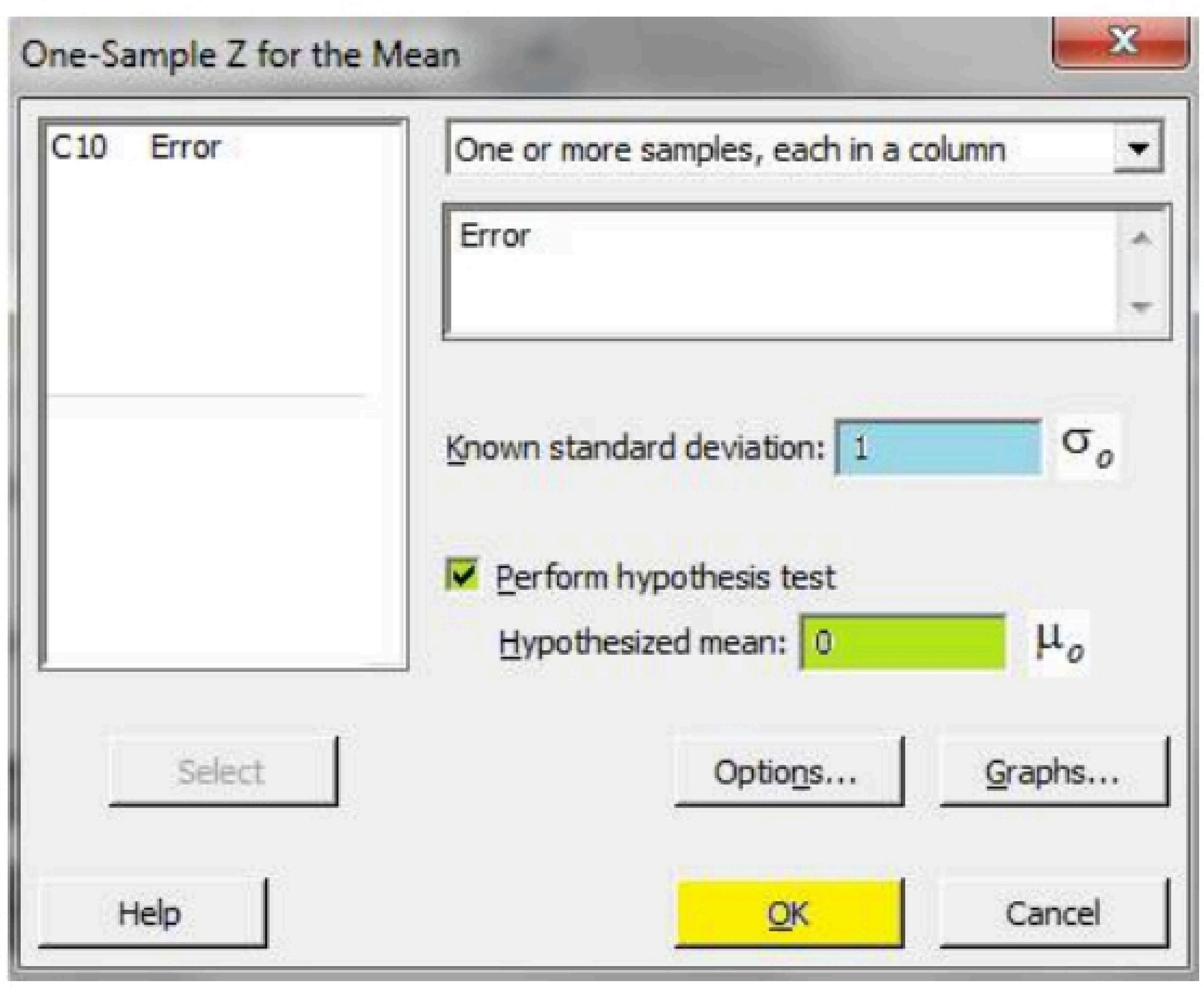
أ – إذا افترضنا أنَّ الانحراف المعياري للمجتمع معلوم ويساوي $\sigma = \sigma_o = 1$ (أي أخذنا بقيمة الانحراف المعياري الذي قدَّمته الشركة)، فعندئذ نتبع الخطوات الآتية (مستخدمين التوزيع الطبيعي):

**Stat → Basic Statistics → 1-Sample Z → Insert Error in free space → For Known standard deviation insert value 1 → For Hypothesized mean insert value 0 → OK.**

يمكنك هنا تغيير مستوى الأهمّية $\alpha$ من أيقونة Options.




ثمَّ الخطوة التالية:




ومن ثمَّ تفعيل Perform hypothesis test مع إدخال المعلومات المطلوبة لتنفيذ الاختبار (قيمتي الانحراف المعياري $\sigma_o$ والمتوسط $\mu_o$ للمجتمع )، وبعد ذلك الضغط على OK فنحصل على التقرير الآتي الذي يحتوي على 95% فترة ثقة للمتوسط $\mu$ بالإضافة لنتيجة الاختبار المنشود.

**One-Sample Z: Error**

**Test of $\mu = 0$ vs $\neq 0$**

**The assumed standard deviation = 1**

| Variable | N | Mean | StDev | SE Mean | 95% CI | Z | P |
|---|---|---|---|---|---|---|---|
| Error | 49 | -0.128 | 1.093 | 0.143 | (-0.408; 0.152) | -0.90 | 0.370 |

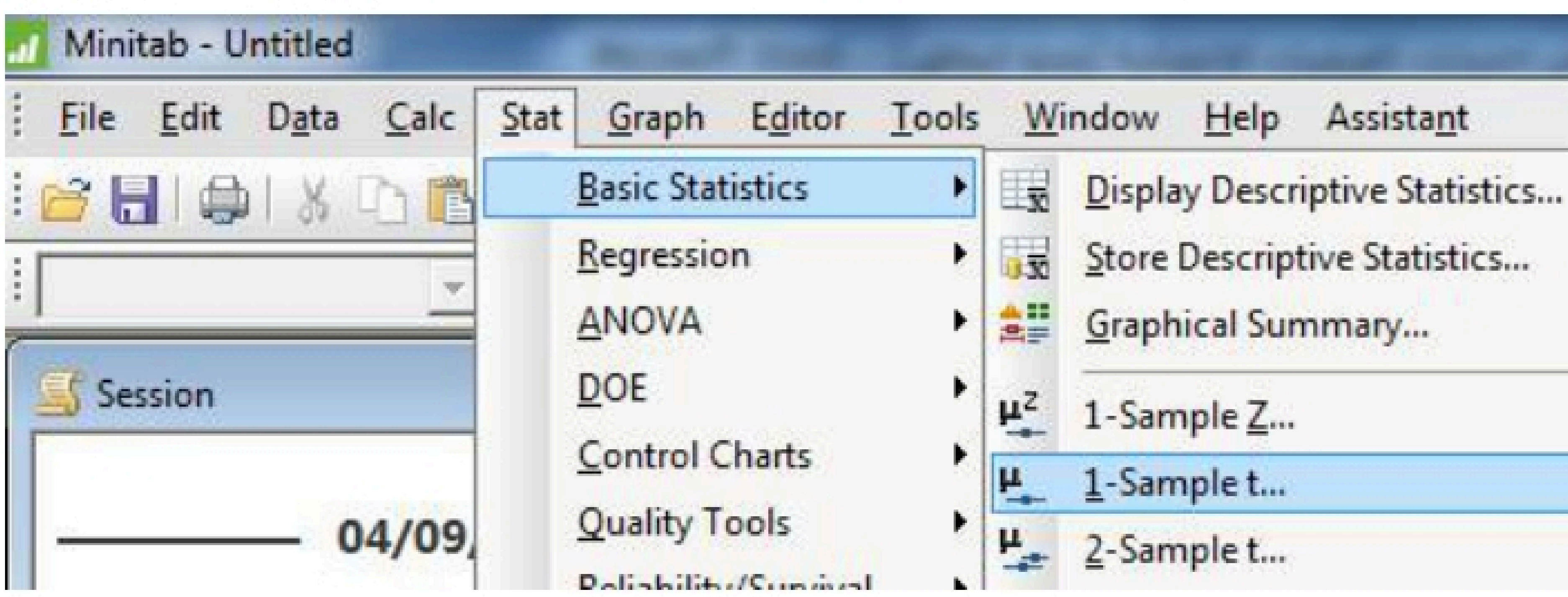
حيث لدينا من هذا التقرير قيمة الإحصاءة $z = -0.90$ ، وكذلك نعلم أنَّ (من جدول التوزيع الطبيعي المعياري) أنَّ $z_{1-\frac{\alpha}{2}} = z_{0.975} = 1.96$ ، ولذلك يكون لدينا:

$$-1.96 < z = \frac{\overline{x} - \mu_o}{\sigma_o}\sqrt{n} = -0.90 < 1.96$$

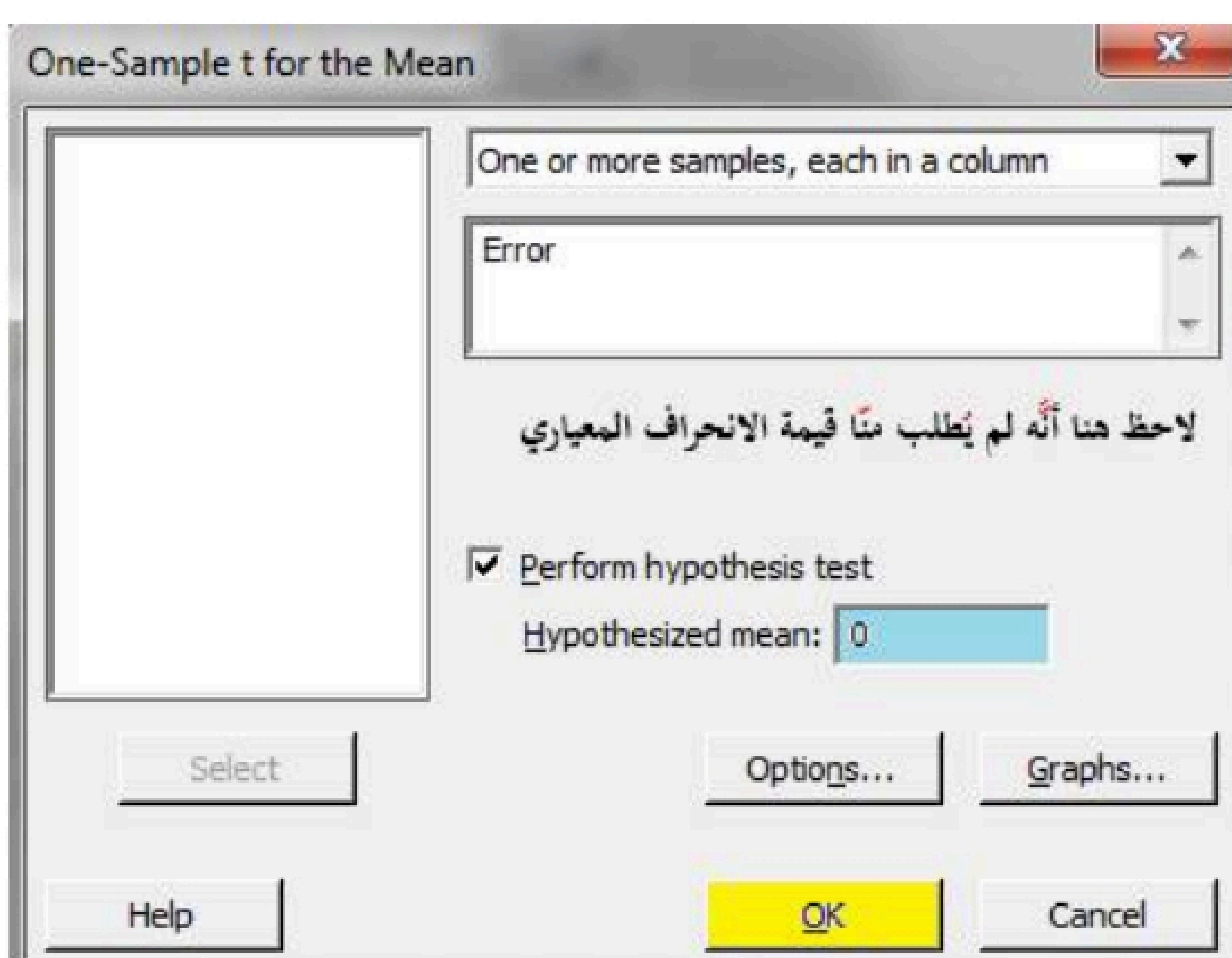
وهذا يعني أنَّه علينا القبول بالفرضية الابتدائية $H_o$ عند مستوى الأهمّيّة $\alpha = 0.05$ . إنَّ هذه النتيجة موجودة في التقرير أيضاً حيث لدينا القيمة المحسوبة **P– value** تساوي 0.370 وهي أكبر من مستوى الأهمّيّة $\alpha = 0.05$ التي تعني قبول الفرضيّة الابتدائية $H_o$ عند مستوى الأهمّيّة $\alpha = 0.05$ .

ب– إذا افترضنا أنَّ الانحراف المعياري للمجتمع مجهولاً $\sigma = ?$ ، فعندئذ نتبع الخطوات الآتية (مستخدمين توزيع **t** - أو ستودنت-):

**Stat ⟶ Basic Statistics ⟶ 1-Sample t ⟶ Known standard deviation ⟶ OK**




ثمَّ الخطوة التالية:




ومن ثمَّ تفعيل Perform hypothesis test مع إدخال المعلومات المطلوبة لتنفيذ الاختبار (قيمة $\mu_o$ فقط)، وبعد ذلك الضغط على OK فنحصل على التقرير الآتي الذي يحتوي على 95% فترة ثقة للمتوسط $\mu$ بالإضافة لنتيجة الاختبار المنشود.

**One-Sample T: Error**
**Test of μ = 0 vs ≠ 0**

| Variable | N | Mean | StDev | SE Mean | 98% CI | T | P |
|---|---|---|---|---|---|---|---|
| Error | 49 | -0.128 | 1.093 | 0.156 | (-0.504; 0.248) | -0.82 | 0.416 |

فنجد من هذا التقرير أنَّ قيمة الإحصاءة $T$ تساوي $-0.82$ ، ولكن لدينا $t^{[n-1]}_{1-\frac{\alpha}{2}} = t^{[48]}_{0.975} = 2.010$ ، ومن ثمَّ يكون لدينا:

$$-2.010 < t = \frac{\bar{x} - \mu_o}{s}\sqrt{n} = -0.82 < 2.010$$

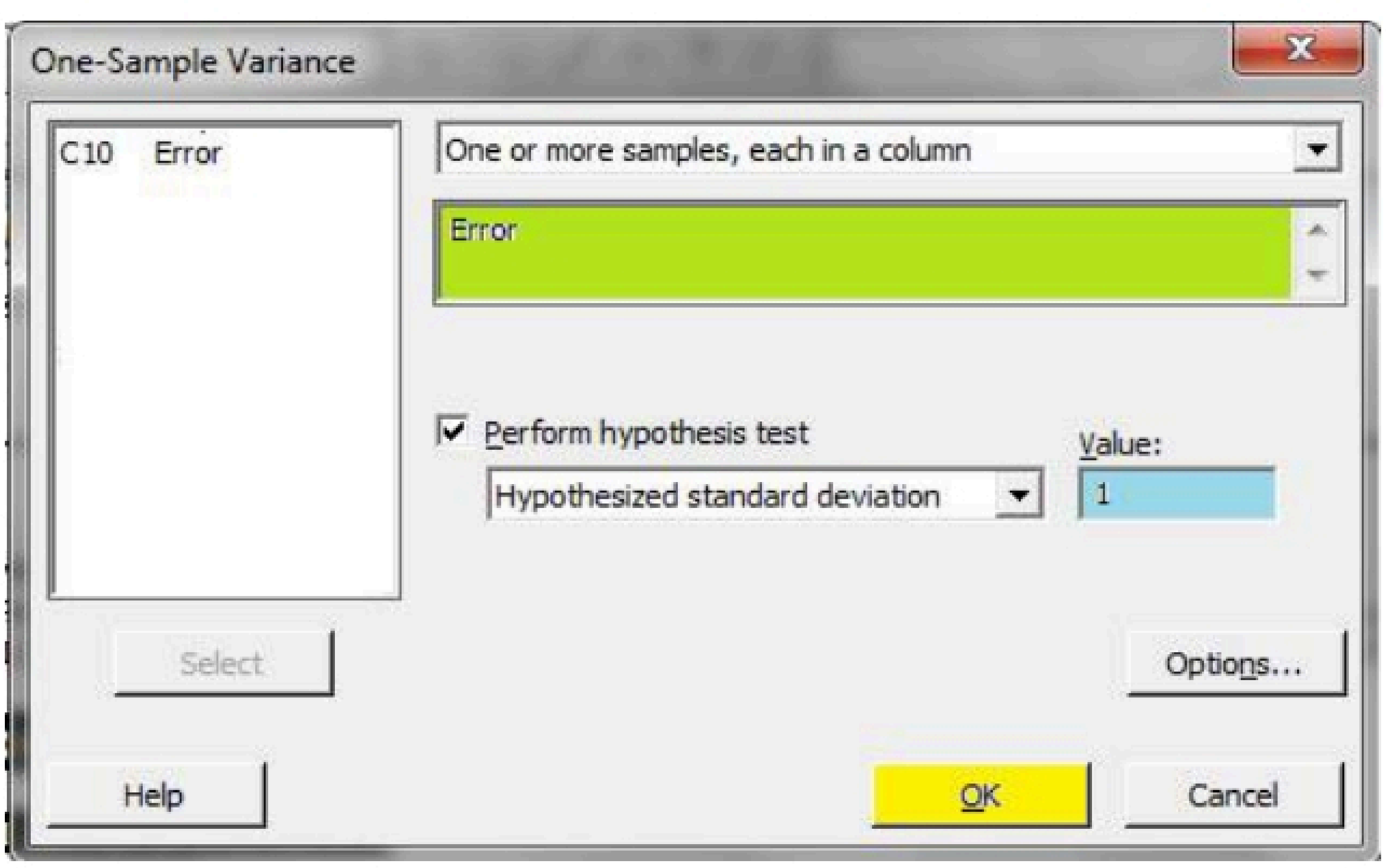
وهذا يعني أنَّه علينا القبول بالفرضية الابتدائية $\mathbf{H}_o$ في هذه الحالة أيضاً، وذلك عند مستوى الأهميّة $\alpha = 0.05$ ، وهذه النتيجة موجودة في التقرير أيضاً حيث لدينا القيمة المحسوبة **P– value** تساوي 0.416 وهي أكبر من مستوى الأهميّة $\alpha = 0.05$ التي تعني قبول الفرضيّة الابتدائية $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha = 0.05$ .

**(١, ١, ٨, ١١) ملاحظة**

إذا كنّا نرغب في تغيير مستوى الأهميّة فإنَّه يمكننا ذلك بالتعديل على هذه القيمة بحسب ما تمَّ شرحه في نهاية الفصل السابق.

الآن، ومن أجل اختبار الفرضيّة الابتدائية القائلة إنَّ تباين مجتمع الأخطاء المرتكبة في نَصْب الأعمدة يساوي الواحد (القيمة التي قدَّمتها الشركة) عند مستوى الأهميّة $\alpha = 0.05$ ، فإنَّه علينا اختبار الفرضيّة الابتدائية $\mathbf{H}_o : \sigma^2 = \sigma_o^2 = 1$ مقابل الفرضيّة البديلة $\mathbf{H}_A : \sigma^2 \neq \sigma_o^2 = 1$ وذلك عند مستوى من الأهميّة قدره $\alpha = 0.05$ ، ولأجل ذلك نتبع الخطوات الآتية:

**Stat ⟶ Basic Statistics ⟶ 1Variance ⟶ Insert Error in free space ⟶ OK.**




فنحصل على التقرير الآتي (تمَّ التعديل على طريقة عرض التقرير) الذي يحتوي نتيجة الاختبار المطلوب بالإضافة إلى 95% فترة ثقة لكل من الانحراف المعياري $\sigma$ والتباين $\sigma^2$ لمجتمع البيانات المقدَّمة.

**Test and CI for One Variance: Error**

**Method**

**Null hypothesis σ = 1 & Alternative hypothesis σ ≠ 1**

**The chi-square method is only for the normal distribution.**

**The Bonett method is for any continuous distribution.**

**Statistics**

| Variable | N | St Dev | Variance |
|---|---|---|---|
| Error | 49 | 1.09 | 1.19 |

**95% Confidence Intervals**

| Variable | Method | CI for St Dev | CI for Variance |
|---|---|---|---|
| Error | Chi-Square | (0.91; 1.37) | (0.83; 1.86) |
| | Bonett | (0.91; 1.37) | (0.82; 1.88) |

**Test**

| Variable | Method | Statistic | DF | P-Value |
|---|---|---|---|---|
| Error | Chi-Square | 57.34 | 48 | 0.335 |

حيث نجد من هذا التقرير أنَّ قيمة الإحصاءة $\chi$ تساوي $\chi = \frac{(n-1)s^2}{\sigma_o^2} = 57.34$ ، ولكن لدينا:

$$\chi_{\frac{\alpha}{2}}^{[n-1]} = \chi_{0.025}^{[48]} \approx \frac{\chi_{0.025}^{[50]} + \chi_{0.025}^{[45]}}{2} = \frac{28.366 + 32.357}{2} = 30.3615$$

$$\chi_{1-\frac{\alpha}{2}}^{[n-1]} = \chi_{0.975}^{[48]} \approx \frac{\chi_{0.975}^{[50]} + \chi_{0.975}^{[45]}}{2} = \frac{65.410 + 71.420}{2} = 68.415$$

ومن ثمَّ يكون $68.415 > \chi > 30.3615$ ، وهذا يعني أنَّه علينا القبول بالفرضية الابتدائية $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha = 0.05$ ، وهذه النتيجة موجودة في التقرير أيضاً حيث لدينا القيمة الاحتمالية **P-value** تساوي 0.335 وهي أكبر من مستوى الأهميّة $\alpha = 0.05$ التي تعني قبول الفرضيّة الابتدائية $\mathbf{H}_o$ عند مستوى الأهميّة $\alpha = 0.05$ .

**(١١,٨,١,٢) ملاحظة**

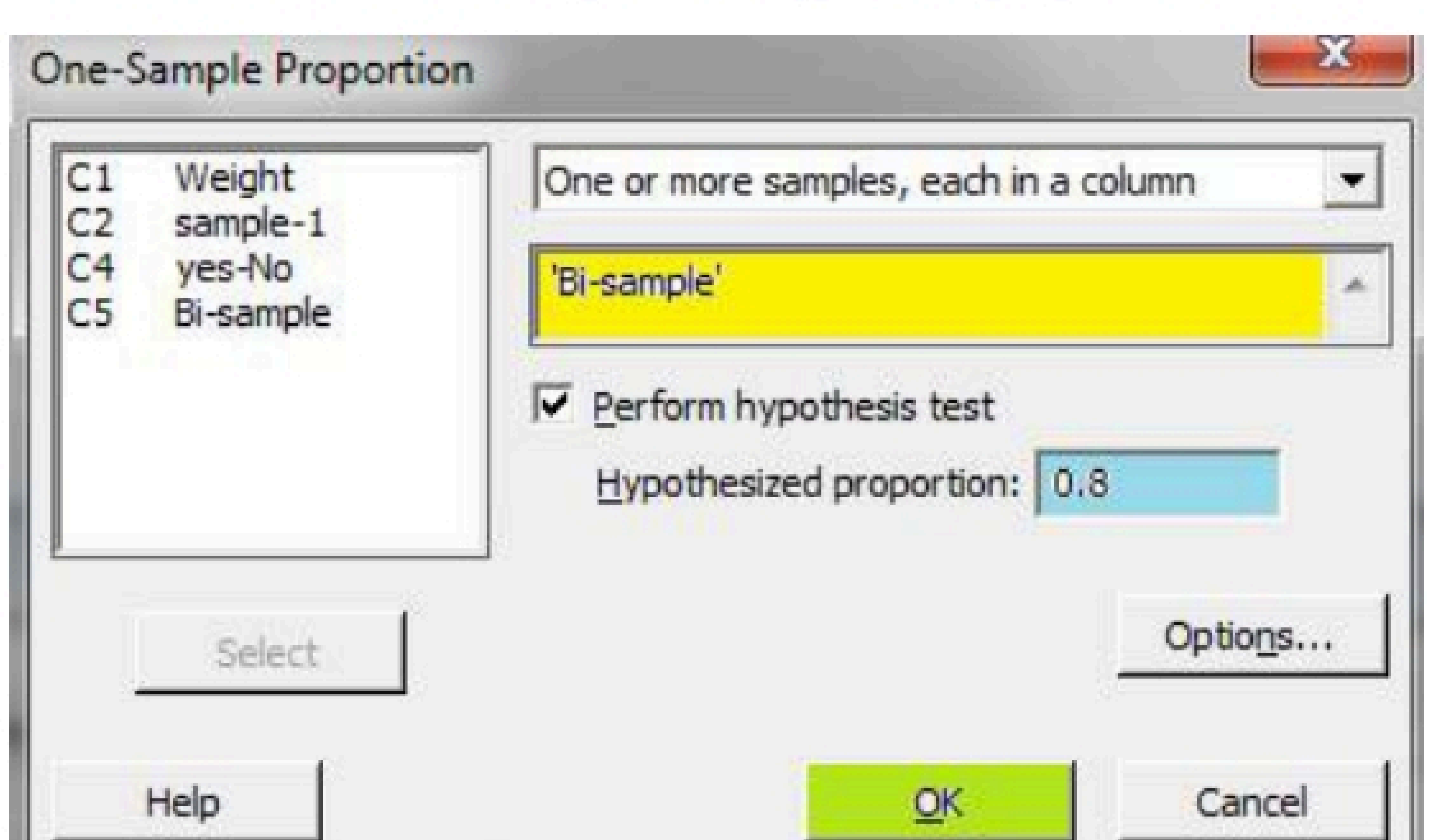
بأسلوب مماثل تماماً يمكننا تنفيذ الاختبارات الإحصائية الخاصّة بفرق المتوسطين ونسبة التباين لمجتمعين طبيعيين.

**(١١,٨,٢) تطبيق**

بالعودة إلى التطبيق (١٠,٤,٢) الذي استخدم في نهاية الفصل السابق، حيث كان لدينا بيانات (تحت مسمًّى **Bi-sample**) تمثِّل عيّنة بحجم 81 شخصاً استطلعت آراؤهم حول مشروع تنموي، علماً أنَّ الإجابة **"موافق"** مُثِّلت بالعدد (1) في حين مُثِّلت الإجابة **"غير موافق"** بالعدد (0) ، ولنختبر الفرضيّة الابتدائية القائلة إنَّ احتمال أن يجيب شخص ما بعبارة " **نعم**" (أي لمَعْلَمة مجتمع آراء الأشخاص $p$) يساوي 0.8 وذلك عند مستوى من الأهميّة قدره $\alpha = 0.05$ ، أي إنَّه علينا اختبار الفرضيّة الابتدائية $\mathbf{H}_o : p = p_o = 0.8$ مقابل الفرضيّة البديلة $\mathbf{H}_A : p \neq p_o = 0.8$ وذلك عند مستوى من الأهميّة قدره $\alpha = 0.05$ .

من أجل ذلك نتبع الخطوات الآتية:

**Stat ⟶ Basic Statistics ⟶ 1-Proportion ⟶ Insert Bi-sample in free space ⟶ Activation Perform hypothesis test ⟶ Insert the value 0.8 in the space for Hypothesis proportion ⟶ OK.**




يمكننا تغيير مستوى الأهميّة $\alpha$ من أيقونة Options.

فنحصل على التقرير الآتي الذي يحتوي على 95% فترة ثقة للمَعْلَمة $p$ بالإضافة لنتيجة الاختبار المنشود والمدوَّنة في السطر الأخير من

التقرير.

**Test and CI for One Proportion: Bi-sample**
**Test of p = 0.8 vs p ≠ 0.8**

| Variable | X | N | Sample p | 95% CI | P-Value |
|---|---|---|---|---|---|
| Bi-sample | 68 | 81 | 0.839506 | (0.741200; 0.911681) | 0.409 |

ومن هذا التقرير نلاحظ أنَّ القيمة الاحتمالية **P – value** تساوي 0.409 وهي أكبر من مستوى الأهمّية $\alpha = 0.05$ ، ومن ثمَّ ليس أمامنا إلاَّ القبول بالفرضية لابتدائية $\mathbf{H}_o$ عند مستوى الأهمّية $\alpha = 0.05$ .

**(١١,٨,٢,١) ملاحظة**

بأسلوب مماثل تماماً يمكننا الحصول على العروض وفترات الثقة الخاصّة بفرق النسبتين $p_1 - p_2$ لمجتمعين برنوليين.

**هذا ما تيسَّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل**

# تمارين الفصل الحادي عشر

١- زعمت إحدى شركات إنتاج الملفات الكهربائية أنَّ متوسط قيم حثِّية الملفات المنتجة لديها يساوي 5000 mH ميلي هنري (**Henry** الهنري وحدة قياس حثيّة الملفات) وبانحراف معياري ? = σ ، قمنا بسحب عيّنة عشوائية من إنتاج هذه الشركة بحجم $n$ فوجدنا أنَّ متوسط قيم حثِّية الملف 4982 ميلي هنري وبانحراف معياري قـدره 28 ميلي هنري، فإذا افترضنا أنَّ القيم الناتجة عن قياسات حثِّية الملفات تتبع التوزيع الطبيعي، فعندئذ:

أ- اختبر الفرضيّة الإحصائية القائلة بصحة ادّعاء الشركة المنتجة عند مستوى من الأهمّية قَدره $\alpha = 0.01$ إذا علمت أنَّ $n = 16$ ملفاً و $\sigma = 25$ ميلي هنري.

ب- اختبر الفرضيّة الإحصائية القائلة بصحة ادّعاء الشركة المنتجة عند مستوى من الأهمّية قَدره $\alpha = 0.01$ إذا علمت أنَّ $n = 16$ ملفاً و $S = 22$ ميلي هنري.

ج- أعد الطلب السابق / أ/ من أجل $n = 81$ ، وماذا تلاحظ؟

د- اختبر الفرضيّة الإحصائية القائلة إنَّ تباين قيم حثِّية الملفات لدى الشركة يساوي 600 عند مستوى من الأهمّية قدره $\alpha = 0.05$ إذا علمت أنَّ $n = 16$ ملفاً، و $S = 22$ ميلي هنري.

هـ- أعد الطلب السابق / د/ من أجل $n = 81$ ، وماذا تلاحظ؟

٢- يَدَّعي منتج لنوع من الدارات الإلكترونية أنَّه على ثقة قدرها 98% بأنَّ أكثر من 20% من الدارات المصنَّعة لديه تعمل بشكل جيد بين درجتي الحرارة 15- و70+ درجة مئوية، وللتأكد من ذلك قمنا بسحب عيّنة حجمها $n$ ، وأُجْريت عليها الدراسة المطلوبة وفقاً لادعاء المنتج، فوجدنا من بينها $m$ دارة تعمل بشكل جيد في نطاق درجتي الحرارة المذكورة آنفاً (علماً أنَّ $n > m$ ).

السؤال هو: هل تؤيد بيانات العيّنة صحة ادّعاء المنتج إذا علمت أنَّه:

أ- لدينا $n = 64$ و $m = 56$ ؟

ب- لدينا $n = 289$ و $m = 253$ ؟

٣- يدَّعي مصنع لإنتاج أسلاك التسخين المستخدمة في المدافئ الكهربائية أنَّ متوسط قيم مقاومة المتر الطولي لهذه الأسلاك يساوي 1.5 أوم ( **Ohm** الأوم وحدة قياس مقاومة مرور التيار الكهربائي في سلك أو ممانعة دارة إلكترونية أو كهربائية، ويرمز لها بـ Ω ) بانحراف قدره σ أوم، وللتأكد من مدى صحة هذا الادعاء قمنا بأخذ عيّنة مكوَّنة من $n$ قطعة طول كل منها يساوي مترًا واحدًا، فوجدنا أنَّ متوسط المقاومة لهذه القطع يساوي 1.6 أوم بانحراف معياري قدره $S$ أوم، فإذا علمت أنَّ قيم المقاومة للمتر الطولي لأسلاك التسخين تتوزع طبيعياً، والمطلوب ما يلي:

أ- إذا كان $\sigma = 0.2$ أوم، فعندئذ اختبر الفرضيّة القائلة إنَّ متوسط قيمة المقاومة للمتر الطولي لهذه الأسلاك أكبر من 1.5 أوم، وذلك عند مسـتوى من الأهمّية قـدره $\alpha = 0.05$ .

ب- إذا كان σ مجهولاً و $S = 1$ أوم و $n = 49$ ، فعندئذ اختبر الفرضيّة القائلة إنَّ متوسط قيمة المقاومة للمتر الطولي لهذه الأسلاك أكبر من 15 أوم وذلك عند مستوى من الأهمّية قدره $\alpha = 0.05$ .

جـ- أعد الطلب السابق إذا كان $n = 256$ .

د- إذا كان $n = 36$ و $S = 1.5$ عندئذ اختبر الفرضيّة القائلة إنَّ قيم المقاومة للمتر الطولي للأسلاك له انحراف معياري قدره 1.25 أوم، وذلك عند مستوى من الأهمّية $\alpha = 0.01$ .

٤- يدّعي مصنع لإنتاج كبسولات لنوع من المضادات الحيوية أنَّ هذه الكبسولات تحوي 500 mg من المادة الفعَّالة بانحراف معياري قدره 5 mg وذلك عند مستوى من الثقة قدره 98% ، وللتأكد من صحة ادّعاء المصنع قمنا بسحب عيّنة حجمها $n$ فوجدنا أنَّ متوسط وزن المادة الفعَّالة في كبسولات العيّنة يساوي 490 mg بانحراف معياري قدره 4 mg ، فعندئذ:

أ- إذا كان حجم العيّنة $n = 25$ كبسولة، فما هو ردك على ادّعاء المصنع مستخدماً اختبار الفرضيّات؟

ب- اختبر الفرضيّة القائلة إنَّ متوسط وزن المادة الفعالة في الكبسولات أقل من 500 mg عند مستوى من الأهميّة $\alpha = 0.05$ ؟

جـ- إذا كان حجم العيّنة $n = 81$ كبسولة، ووجدنا أنَّ 70 كبسولة منها مُحقَّقةٌ للمواصفات المقدَّمة من قبل المصنع، فعندئذ اختبر الفرضيّة القائلة إنَّ نسب الكبسولات المنتجة والمحقَّقة للمواصفات المقدَّمة من قبل المصنع يساوي 88% وذلك عند مستوى من الأهميّة قدره $\alpha = 0.001$ .

٥- ادعت وزارة البيئة في بلد ما أنَّها على ثقة قدرها 95% بأنَّه أكثر من 85% من المصانع الموجودة في البلد تُحقِّق معايير مكافحة تلوث البيئة، ولكن أنَّصار المحافظة على البيئة أرادوا التَحقّق من صحة هذا الادعاء، فقاموا بدراسة ميدانية للكشف عن التلوث (وفقاً لمعايير متفق عليها) في 64 موقعاً اختيرت بشكل عشوائي، فوجـدوا أنَّ 56 موقعاً يُحقِّق معايير مكافحة التلوث، والمطلوب:

أ- ما هو رد أنَّصار المحافظة على البيئة إذا كان اختبارهم مبنياً على مستوى الأهميّة $\alpha = 0.05$ ؟

ب- ما هو حجم العيّنة $n$ بحيث إنه عند مستوى من الأهميّة $\alpha = 0.01$ يكون رد أنَّصار المحافظين على البيئة إيجابياً إذا كانت النسبة المستحصل عليها من العيّنة هي 80%؟

جـ- إذا كان $n = 64$ موقعاً، ووجد أنَّ 53 موقعاً تُحقِّق معايير مكافحة التلوث للبيئة، فعندئذ أي مسـتوى من الأهميّة يكون فيه رد أنَّصار المحافظين على البيئة إيجابياً؟

٦- في مصنع لإنتاج القضبان الحديدية يوجد خطان لإنتاج القضبان ذات القطر 6 مم وطول 6 م، ويعملان بشكل مستقل كل منهما عن الآخر، فقمنا بسحب عيّنة $\mathcal{X}_{36}$ و $\mathcal{Y}_{20}$ من إنتاج الخط الأول والثاني على الترتيب، فوجدنا أنَّ متوسط قطر القضبان للعيِّنة $\mathcal{X}_{36}$ و $\mathcal{Y}_{20}$ هي 6.15 مم و5.95 مم على الترتيب وبانحراف معياري على الترتيب هو 0.25 مم و0.35 مم على الترتيب أيضاً. الآن، وبفرض أنَّ قيم قياسـات أقطار القضبان تخضع للتوزيـع الطبيعي، فعنـدئذ ا ختبر الفرضيّة الإحصائية القائلة:

أ- إنَّ متوسط أقطار القضبان المنتـجة من الخطين متساوية عند مستوى من الأهميّة قدره $\alpha = 0.05$ .

ب- إنَّ تباين قيم الأقطار للقضبان المنتجة من الخطين متساوية عند مستوى من الأهميّة قَدره $\alpha = 0.01$ .

٧- قام رجلان بالرمي على هدف واحد، وبشكل مستقل كل منهما عن الآخر، وبحيث لا يعلم أي منهما شيئاً عن إصابته للهدف أثناء عملية الرمي، فإذا قام الأول بإطلاق 180 رمية على الهدف تحقَّق منها 170 إصابة للهدف، في حين أطلق الثاني 150 رمية على الهدف تحقَّق منها 143 إصابة للهدف، فهل يمكننا الادعاء أنَّ للرجلين الكفاءة نفسها في الرمي عند مستوى من الأهميّة $\alpha = 0.05$ ، وهل يبقى القرار نفسه من أجل $\alpha = 0.45$ ؟

٨- لدى رصد عدد الحوادث المرورية على طريق عام معيَّن وخلال الفترة من الساعة 14 إلى 15 (2 وحتى 3 بعد الظهر) وجـدت البيانات الآتية لـ 16 يوماً متفرقة اختيرت عشوائياً من عام مُحدَّد:

| اليوم | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| $X$ عدد الحوادث | 3 | 0 | 5 | 4 | 1 | 0 | 0 | 2 |
| اليوم | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 |
| $X$ عدد الحوادث | 1 | 1 | 2 | 0 | 3 | 0 | 0 | 1 |

والمطلوب اختبر الفرضيّة القائلة إنَّ عدد الحوادث وفقاً لما ذكر يخضع للتوزيع البواسوني عند مستوى من الأهميّة $\alpha = 0.01$ .

٩- أخذت عيّنة بحجم 12 عداءً من مجتمع المتقدمين لسباق المارثون فوجد زمن وصولهم للهدف المنشود مُعطى كما في الجدول الآتي:

| رقم العدّاء | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| $X$ الزمن المستغرق | 2:42 | 2:50 | 2:55 | 2:40 | 3:20 | 2:54 |
| رقم العدّاء | 7 | 8 | 9 | 10 | 11 | 12 |
| $X$ الزمن المستغرق | 3:15 | 3:00 | 3:15 | 3:50 | 3:10 | 3:00 |

والمطلوب اختبر الفرضيّة القائلة إنَّ زمن وصول المتسابقين للهدف يخضع للتوزيع الأسي وذلك عند مستوى من الأهميّة $\alpha = 0.05$ .

١٠- في مركز لاختبار العيّنات للبضائع الصناعية والزراعية يوجد ثلاثة معامل مخبرية $L_1$، $L_2$ و $L_3$ تقوم بهذه الاختبارات، ولدى معاينة عيّنة من نتائج المخبر $L_1$، $L_2$ و $L_3$ بحجم $n_1 = 49$، $n_2 = 64$ و $n_3 = 81$ على الترتيب، لاحظنا أنّه يوجد لدى بعض هذه المعامل المخبرية نتائج غير صحيحة وعددها مُعطى كما في الجدول الآتي:

| رقم العدّاء | $L_1$ | $L_2$ | $L_3$ |
|---|---|---|---|
| نتائج صحيحة | 67 | 41 | 89 |
| رقم العدّاء | 7 | 8 | 9 |
| نتائج خاطئة | 4 | 3 | 5 |

والمطلوب اختبار الفرضيّة القائلة إنَّ نسب النتائج الصحيحة المقدَّمة من قبل هذه المعامل المخبرية متساوية عند مستوى من الأهميّة قدره $\alpha = 0.05$ .

١١- لدى أخذ عيّنة بحجم 86 قضيب حديدي بطول متر واحد وقطر 10 مم من إنتاج مصنع للحديد والصلب وإخضاع هذه العيّنة لفحص قوة الشد مقدرة بـ 10 ميلي بار، وكانت نتائج تحديد نسبة الكربون الموجودة في معدن القضيب مقدَّرة بالنسبة المئوية كما في الجدول الآتي:

| | | قوة الشدّ | | | | | | المجموع |
|---|---|---|---|---|---|---|---|---|
| | | 60→64 | 64→68 | 68→72 | 72→76 | 76→80 | 80→84 | |
| نسبة الكربون | 0.35→0.37 | 1 | 1 | 0 | 0 | 0 | 0 | |
| | 0.37→0.39 | 1 | 3 | 2 | 1 | 0 | 0 | |
| | 0.39→0.41 | 2 | 3 | 7 | 3 | 0 | 0 | |
| | 0.41→0.43 | 1 | 4 | 11 | 12 | 1 | 0 | |
| | 0.43→0.45 | 1 | 1 | 3 | 7 | 5 | 0 | |
| | 0.45→0.47 | 0 | 0 | 4 | 3 | 5 | 1 | |
| | 0.47→0.49 | 0 | 0 | 0 | 0 | 3 | 0 | |
| Total | | | | | | | | |

والمطلوب اختبا ر الفرضيّة القائلة إنَّ قوة تحمُّل القضيب للشدِّ مستقلّة عن نسبة الكربون الموجودة في معدن القضيب وذلك عند مستوى الأهميّة $\alpha = 0.01$ .

١٢- اختيرت عينتان عشوائيتان من طلبة قسم الكيمياء في جامعتين $X$ و $Y$ للقيام بتنفيذ تجربة معيّنة وتحت الشروط نفسها، وبحيث إن المهارات الفردية تقوم بدور فعّال في تحقيق النتائج المرجوة من هذه التجربة، فإذا علمنا أنَّ كل طالب يقوم بإجراء التجربة بشكل مستقل عن الآخرين وأنَّ حجم العيّنة الأولى (من الجامعة X) يساوي $n$، وحجم العيّنة الثانية (من الجامعة Y) يساوي $m$، وبفرض أنَّ التجربة تحقَّقت بشكل إيجابي من قبل $n_1$ و $m_1$ طالب في العيّنة الأولى والثانية على الترتيب. عندئذ اختبر الفرضيّة القائلة بتساوي الكفاءة لدى طلبة الجامعتين عند مستوى من الأهميّة قدره $\alpha$، وذلك من أجل الحالات الآتية:

أ- لدينا $n = 36$ ، $n_1 = 25$ ، $m = 49$ ، $m_1 = 39$ و $\alpha = 0.05$ .

ب- لدينا $n = 360$ ، $n_1 = 250$ ، $m = 490$ ، $m_1 = 390$ و $\alpha = 0.05$ .

ج- لدينا $n = 36$ ، $n_1 = 25$ ، $m = 49$ ، $m_1 = 39$ و $\alpha = 0.01$ .

١٣- في مصنع لإنتاج المصابيح الكهربائية يوجد خمسة خطوط للإنتاج $L_1$ ، $L_2$ ، $L_3$ ، $L_4$ و $L_5$ ويعمل كل منها بشكل مستقل عن الخطوط الأخرى، وأنَّ كل خط ينتج ثلاثة أنواع من المصابيح ذات الاستطاعة 60 و90 و150 واط. الآن إذا افترضنا أنَّ عمليّة الإنتاج لهذه الخطوط تتبع توزيعاً طبيعياً بمتوسط $\mu_1$ ، $\mu_2$ ، $\mu_3$ ، $\mu_4$ و $\mu_5$ وبانحراف معياري $\sigma_1$ ، $\sigma_2$ ، $\sigma_3$ ، $\sigma_4$ و $\sigma_5$ على الترتيب علماً أنَّ $\sigma_1 = \sigma_2 = \sigma_3 = \sigma_4 = \sigma_5$ ، وقمنا بسحب عيّنات عشوائية من إنتاج هذه الخطوط بحجم:

$$n_1 = 255 \ \& \ n_2 = 255 \ \& \ n_3 = 242 \ \& \ n_4 = 245 \ \& \ n_5 = 248$$

على الترتيب، وكانت نتائج دراسة هذه العيّنات كما في الجدول الآتي:

| | | الاستطاعة | | | المجموع |
|---|---|---|---|---|---|
| | | 60W | 90W | 150 W | |
| الخط | $L_1$ | 87 | 78 | 90 | |
| | $L_2$ | 83 | 81 | 91 | |
| | $L_3$ | 79 | 79 | 84 | |
| | $L_4$ | 81 | 82 | 82 | |
| | $L_5$ | 80 | 80 | 88 | |
| **Total** | | | | | |

عندئذ اختبر الفرضيّة القائلة بتساوي متوسطات الإنتاج لهذه الخطوط الخمسة عند مستوى الأهميّة $\alpha = 0.05$ .

١٤- في مستشفى يوجد خمس عيادات $B_1$ ، $B_2$ ، $B_3$ ، $B_4$ و $B_5$ تعمل كل منها بشكل مستقل عن الأخرى، قمنا بسحب عيّنات عشوائية من سجلات المرضى الذين زاروا هذه العيادات بحجم $n_1 = 85$ ، $n_2 = 77$ ، $n_3 = 95$ ، $n_4 = 118$ و $n_5 = 95$ على الترتيب فوجد أنَّ عدد المرضى الذين تماثلوا للشفاء في هذه العيادات مقدَّم كما في الجدول الآتي:

| | | العيادة | | | | | المجموع |
|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | |
| الحالة | تماثل للشفاء | 82 | 75 | 93 | 105 | 88 | |
| | لم يتماثل للشفاء | 3 | 2 | 2 | 13 | 7 | |
| **Total** | | | | | | | |

والمطلوب اختبار الفرضيّة القائلة بتساوي نسب المرضى الذين تماثلوا للشفاء في هذه العيادات الخمس وذلك عند مستوى الأهميّة $\alpha = 0.01$ .

# الفصل الثاني عشر

## المتسلسلات الزمنية

## TIME SERIES

### (١٢,١) مفاهيم أولية في المتسلسلات الزمنية

### Initial Concepts in Time Series

لقد قدّمنا في الفصول الثلاث الأولى من هذا الكتاب دراسات أولية متنوعة حول البيانات وتحليلها، وذكرنا في فصل الارتباط و تحليل الانحدار أنَّه من الممكن أن يكون لدينا بيانات متزاوجة $(x_1,y_1)$، $(x_2,y_2)$،...و $(x_n,y_n)$ فيها البيانات $x_1$، $x_2$،... و $x_n$ هي لحظات زمنية متتالية (وسنرمز لها فيما بعد بـ $t_1$، $t_2$،... و $t_n$)، وأمّا القيم $y_1$، $y_2$،... و $y_n$ فهي قيم مقيسه عند هذه اللحظات الزمنية على الترتيب (وسنرمز لها بـ $x_{t_1}$، $x_{t_2}$،...و $x_{t_n}$). أكثر من ذلك، فقد تكون لدينا عدّة قيم مقيسه $x_{1i}$، $x_{2i}$،...و $x_{mi}$ موافقة للحظة زمنية $t_i$ من أجل أي $i \in \mathbb{N}_n$، والأمثلة البسيطة الآتية توضِّح لنا بعضاً منها:

١ – القيمة المسجَّلة كل ساعة لأسعار أسهم شركة ما في البورصة.

٢ – كميّة الهطل المطري الشهري على منطقة محدَّدة.

٣ – القياسات المأخوذة لدرجة الحرارة والضغط لمريض كل ساعتين في مستشفى.

٤ – قيمة المبيعات اليومية من اللحوم، الخضراوات، الملابس والأجهزة الكهربائية في سوق تجاري.

٥ – عدد القتلى في معارك لبلد ما مع مرور الزمن.

٦ – عدد حوادث الطيران المدني في عام 2014.

فنلاحظ من الأمثلة السابقة أنَّ البيانات التي يمكن أن نحصل عليها مرتبطة بالزمن. إنَّ دراسة وتحليل هذا النوع من البيانات يتطلب منهجاً خاصاً في كثير من جوانبه، ولذلك تُصنَّف هذه البيانات في مُسمَّى خاص بها يُدعى المتسلسلات الزمنية والذي يقدَّمه لنا التعريف الآتي.

#### (١٢ , ١ , ١) تعريف (المتسلسلة الزمنية Time Series)

المتسلسلة زمنية هي مجموعة من المشاهدات (أو الملحوظات لأشياء معرَّفة جيداً) تمَّ الحصول عليها من خلال القياسات المتكرّرة على مرّ الزمن. بعبارة موجزة: المتسلسلة الزمنية هي مجموعة بيانات مُرتّبة زمنياً.

الآن وبناءً على هذا التعريف، وبإمعان النظر في الأمثلة المذكورة سابقاً فإنَّنا نلاحظ ما يلي:

١ – بعد الأخذ بالحسبان أنَّ الزمن جزء لا يتجزأ من مفهوم المتسلسلة الزمنية، فإنَّ المتسلسلة الزمنية يمكن أن يكون لها أكثر من بُعد واحد وذلك بحسب عدد القيم المأخوذة عند كل لحظة زمنية، فعلى سبيل المثال

أ – نجد المثالين الأول والثاني السابقين يقدّمان لنا متسلسلتين زمنيتين لهما بُعد يساوي الواحد، وفي دراستنا المقبلة لن نقدِّم إلاَّ هذا النوع من المتسلسلات الزمنية.

ب- يقدّم لنا المثال الثالث متسلسلة زمنية بُعدها يساوي 2، وكذلك المثال الرابع يقدّم لنا متسلسلة زمنية بُعدها يساوي 4، وهذه الأنواع من المتسلسلات الزمنية ستكون خارج إطار دراستنا في هذا الكتاب.

٢- من الممكن أن تؤخذ القياسات على فترات زمنية (تُدعى فجوات زمنية **Lags**) غير منتظمة (أي على فترات زمنية غير متساوية الطول)، والمثال الخامس يقدّم لنا متسلسلة زمنية من هذا النوع؛ وذلك لأنَّ المعارك التي تحصل لبلد ما لا تحدث على فترات زمنية متساوية الطول في الحالة العامّة، وهذا النوع من المتسلسلات الزمنية لن نتناوله في دراساتنا المقبلة أيضاً. أي إننا في دراستنا المقبلة سندرس متسلسلات زمنية قيمها تحدث على فترات زمنية متساوية الطول (لها فجوات زمنية متساوية الطول).

٣- البيانات التي يتمّ جمعها مرّة واحدة فقط ليست متسلسلة زمنية، ومن ثمَّ المثال السادس لا يقدّم لنا متسلسلة زمنية.

في الواقع إنَّ دراسة المتسلسلات الزمنية يمكن أن تقدَّم على مستويات عديدة من المعرفة العلميّة، فيمكن أن تقدَّم بشكلها التقليدي البسيط، أو تقدَّم بعروض رياضياتية متقدَّمة مبنية على دراسات متقدَّمة في نظرية الطوريات العشوائية ونظرية تحليل الانحدار، وقد تقدَّم في مستوى وسط بين المستويين السابقين. في هذا الفصل سوف نعتمد على المستوى الأول ونقدَّم المعلومات بشكلها التقليدي البسيط من أجل التعريف بهذا العلم وفهم المصطلحات الأساسيّة التي تتعامل معها. علماً أنَّنا سنعتمد وبشكل أساس على المراجع [12]، [33] و[67] في تقديم هذا الفصل، ونوجِّه القارئ إلى المراجع الآتية (الموجودة في فهرس مراجع هذا الكتاب) إذا رغب في معرفة المزيد حول هذا النوع من الدراسات العشوائية التطبيقية [9]،[11]،[21]،[25]،[32]،[34] و[48].

من المفاهيم الضرورية في إطار دراسة المتسلسلات الزمنية ما يُعرف باسم الطوري العشوائي الذي سنقدِّم تعريفه على النحو الآتي.

### (١٢ ,١ ,٢) تعريف (الطوري العشوائي الحقيقي Real Random Process)

ليكن $[\Omega,\mathscr{A},P]$ فضاءً احتمالياً، و $T\subseteq\mathbb{R}$ مجموعة غير خالية، ولنأخذ $X: T\times\Omega\to\mathbb{R}$ بحيث $(t,\omega)\mapsto X(t,\omega)$ تطبيقاً قيوساً بالنسبة إلى $\mathscr{A}$ و $\Re$ من أجل كل قيمة $t\in T$ مثبّتة، فعندئذ أسرة كل المتغيِّرات العشوائية $\{X(t,\bullet);t\in T\}$ تُدعى طورياً عشوائياً حقيقياً (أو طورياً عشوائياً **Random Process** حيث سنقتصر دراستنا هذه على الطوريات العشوائية الحقيقية فقط).

### (١٢ ,١ ,٢ ,١) ملاحظات

١- يمكن تعريف الطوري العشوائي بطريقة أخرى أيضاً وذلك من خلال أخذ $T\subseteq\mathbb{R}$ مجموعة غير خالية، وكذلك أخذ $X_t:\Omega\to\mathbb{R}$ بحيث $\omega\mapsto X_t(\omega)$ تطبيقاً قيوساً بالنسبة إلى $\mathscr{A}$ و $\Re$ من أجل كل $t\in T$ مثبّتة، فعندئذ تُدعى أسرة كل المتغيِّرات العشوائية $\{X_t;t\in T\}$ طورياً عشوائياً (ويُكتب بالشكل الآتي أيضاً $\{X_t\}_{t\in T}$).

٢- من أجل أيّ حادث ابتدائي $\omega\in\Omega$ مثَّبت يكون $X_\bullet(\omega)$ دالَّة حقيقية على $T$، وفي هذه الحالة يُطلق على الدالَّة $X_t(\omega)$ (وكذلك رسمها البياني كدالَّة بـ $t$) اسم مسار Trajectory للطوري العشوائي $\{X_t;t\in T\}$، ومن ثمَّ يكون للطوري العشوائي $\{X_t;t\in T\}$ مسارات عددها يساوي قدرة فضاء الحوادث الابتدائية $\Omega$.

٣- في كثير من الدراسات يتمّ التعامل مع المجموعة $T$ (والتي تُدعى مجموعة مَعَالم الطوري العشوائي) على أنَّها فترة زمنية، ولذلك ينظر إلى القيم $t\in T$ على أنَّها لحظات زمنية، وسنعتمد هذا المفهوم في دراستنا للمتسلسلات الزمنية.

٤- إذا كانت مجموعة المَعَالم $T=\mathbb{N}_n$ مع $n\in\mathbb{N}$ مثبتة، فعندئذ يتحدث المرء عن متتالية منتهية من المتغيِّرات العشوائية أو متَّجه عشوائي حقيقي.

٥- إذا كانت مجموعة المَعالِم $T$ قابلة للعدِّ على الأكثر، فعندئذ يتحدث المرء عن متتالية من المتغيِّرات العشوائية (وفي مثل هذه الحالة غالباً ما تؤخذ $T \supseteq \mathbb{N}^o$).

٦- لكلّ متسلسلة زمنية يوجد طوري عشوائي $\{X_t \, ; \, t \in T\}$ يولِّدها، ولذلك فمن أجل عيِّنة محدَّدة من مسار هذه المتسلسلة الزمنية سنرمز لقيمة هذا الطوري العشوائي (قيمة إحدى مشاهداته) عند لحظة زمنية $t \ni T$ بالرمز $x_t$.

٧- بفرض أنَّه لدينا $(t_1, x_{t_1})$، $(t_2, x_{t_2})$،...و $(t_\tau, x_{t_\tau})$ عيِّنة حجمها يساوي $\tau$ مأخوذة من مسار متسلسلة زمنية، وننوِّه هنا إلى أنَّ استخدام الحرف $\tau$ دارج في مجال دراسة المتسلسلات الزمنية للتعبير عن حجم البيانات المأخوذة من مسار متسلسلة زمنية، ولكن تستخدم الرموز المألوفة سابقاً $n$ أو $m$ بين الحين والآخر أيضاً.

٨- بما أنَّ دراستنا المقبلة مركَّزة على المتسلسلات الزمنية التي قيمها تحدث على فترات زمنية متساوية الطول، فإنَّنا سنقوم (وفي كثير من الحالات) بإلباس اللحظة الزمنية بدليلها على سبيل التبسيط. أي سنكتب 1 عوضاً عن $t_1$ وكذلك 2 عوضاً عن $t_2$ وهكذا دواليك حتى اللحظة الزمنية الأخيرة $t_\tau$ سنكتبها $\tau$، ومن ثَّم يصبح لدينا العرض الآتي للمشاهدات السابقة $(1, x_1)$، $(2, x_2)$،...و $(\tau, x_\tau)$ (هذا على سبيل التبسيط فقط).

٩- لقد استخدمنا هنا الرمز $\tau$ لكي نشير به إلى أنَّ القيمة $x_\tau$ وهي آخر قيمة تمَّ الحصول عليها، ومن ثمَّ فهي تشير إلى القيمة التي حصلنا عليها في الوقت الحاضر، وما يأتي بعدها من قيم مثل $x_{\tau+1}$، $x_{\tau+2}$ و... إنّما هي قيم متنبأ بها للمستقبل، وأمّا القيم السّابقة لـ $x_\tau$ مثل $x_{\tau-1}$، $x_{\tau-2}$،...، $x_2$ و $x_1$ فهي قيم من ماضي المتسلسلة الزمنية وأقدمها $x_1$ التي تمثِّل أو قيمة تمَّ الحصول عليها في العينة.

الآن، ومن أجل دراسة متسلسلة زمنية مُعطاة نقوم عادة بتجزئة هذه الدراسة على مراحل عدّة، وأول هذه المراحل هي تقديم ما يُعرف باسم بصمة المتسلسلة الزمنية (أو العرض الانتشاري لمتسلسلة زمنية)، حيث يمكن لهذا العرض أن يقدِّم لنا تصوراً أولياً حول مسار وسلوك المتسلسلة الزمنية. إنَّ هذا المفهوم يقدِّمه لنا التعريف الآتي.

**(١٢,١,٣) تعريف (بصمة متسلسلة زمنية Imprint of Time Series)**

لتكن $x_1$، $x_2$،...و $x_\tau$ مشاهدات مأخوذة من مسار متسلسلة زمنية، فلو قمنا بالتمثيل النقطي لهذه المشاهدات على لوحة الانتشار كما سبق وقدمناه في فصل الارتباط والانحدار، ومن ثمَّ الوصل بين هذه النقاط على التتالي بقطع مستقيمة، ومع الافتراض أنَّ المحور الأفقي هو محور الزمن $T$ حيث تدوَّن عليه اللحظات الزمنية، في حين أنَّ المحور الرأسي هو محور القيم وتعيّن عليه القيم $x_1$، $x_2$،...و $x_\tau$، فعندئذ يُعرف الشكل البياني النّاتج باسم **بصمة المتسلسلة الزمنية**.

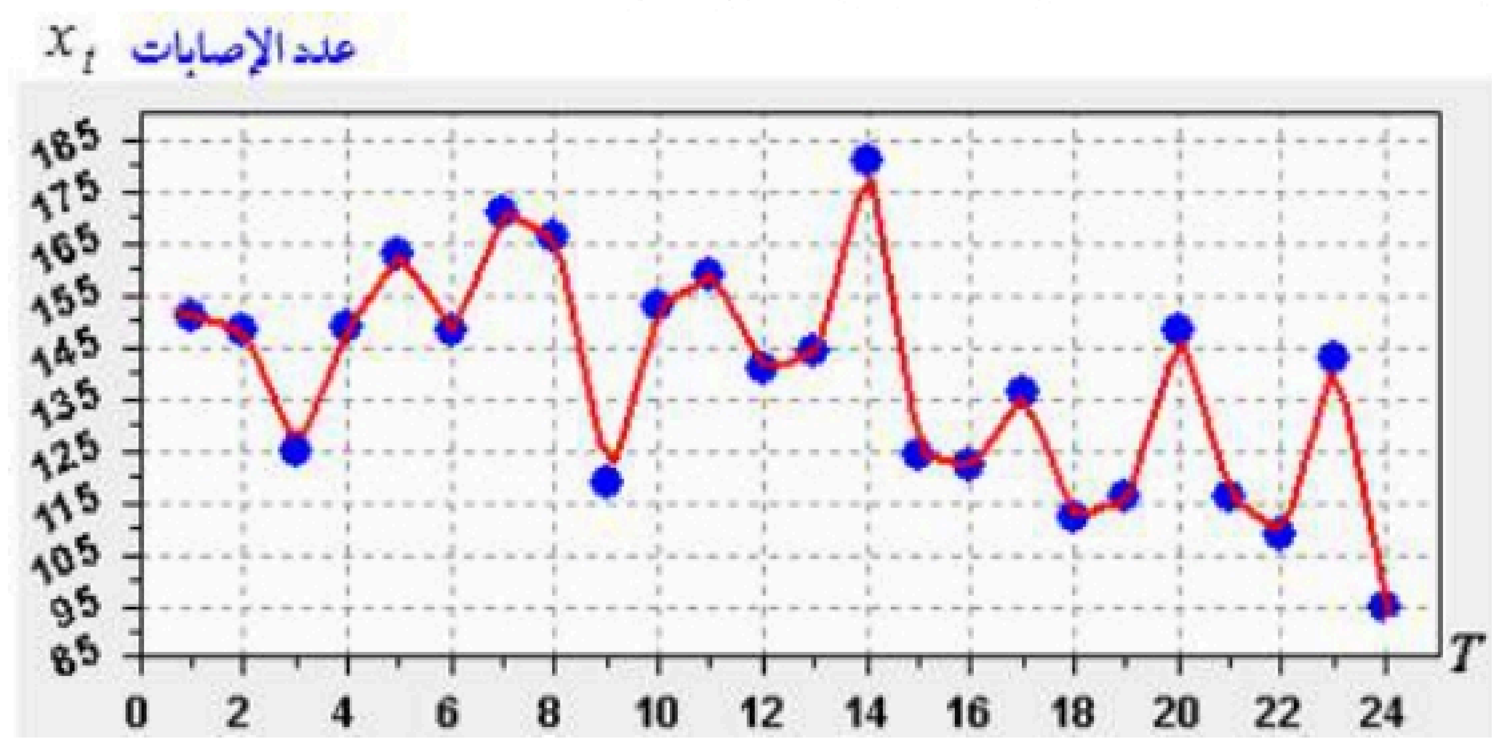
**(١٢,١,٣,١) أمثلة**

١- لتكن لدينا المشاهدات الآتية التي تمثّل عدد الإصابات غير المميتة في حوادث مرورية في مدينة الرياض خلال العامين 1423 و1424هـ.

الجدول (١٢,١)

| العام هـ | الشهر | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| 1423 | عدد الإصابات | 151 | 148 | 125 | 149 | 163 | 148 |
| 1424 | عدد الإصابات | 144 | 181 | 124 | 122 | 136 | 112 |
| العام هـ | الشهر | 7 | 8 | 9 | 10 | 11 | 12 |
| 1423 | عدد الإصابات | 171 | 166 | 119 | 153 | 159 | 141 |
| 1424 | عدد الإصابات | 116 | 148 | 116 | 109 | 143 | 95 |

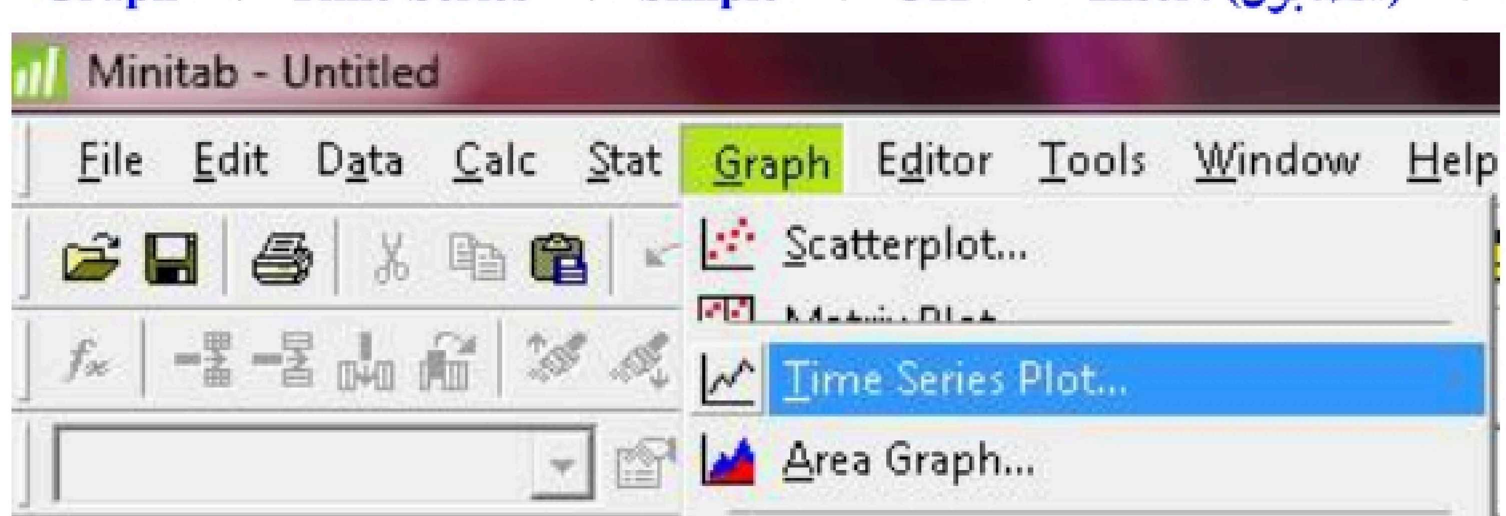
فعندئذ سيكون لبصمة هذه المتسلسلة الزمنية العرض البياني الآتي:



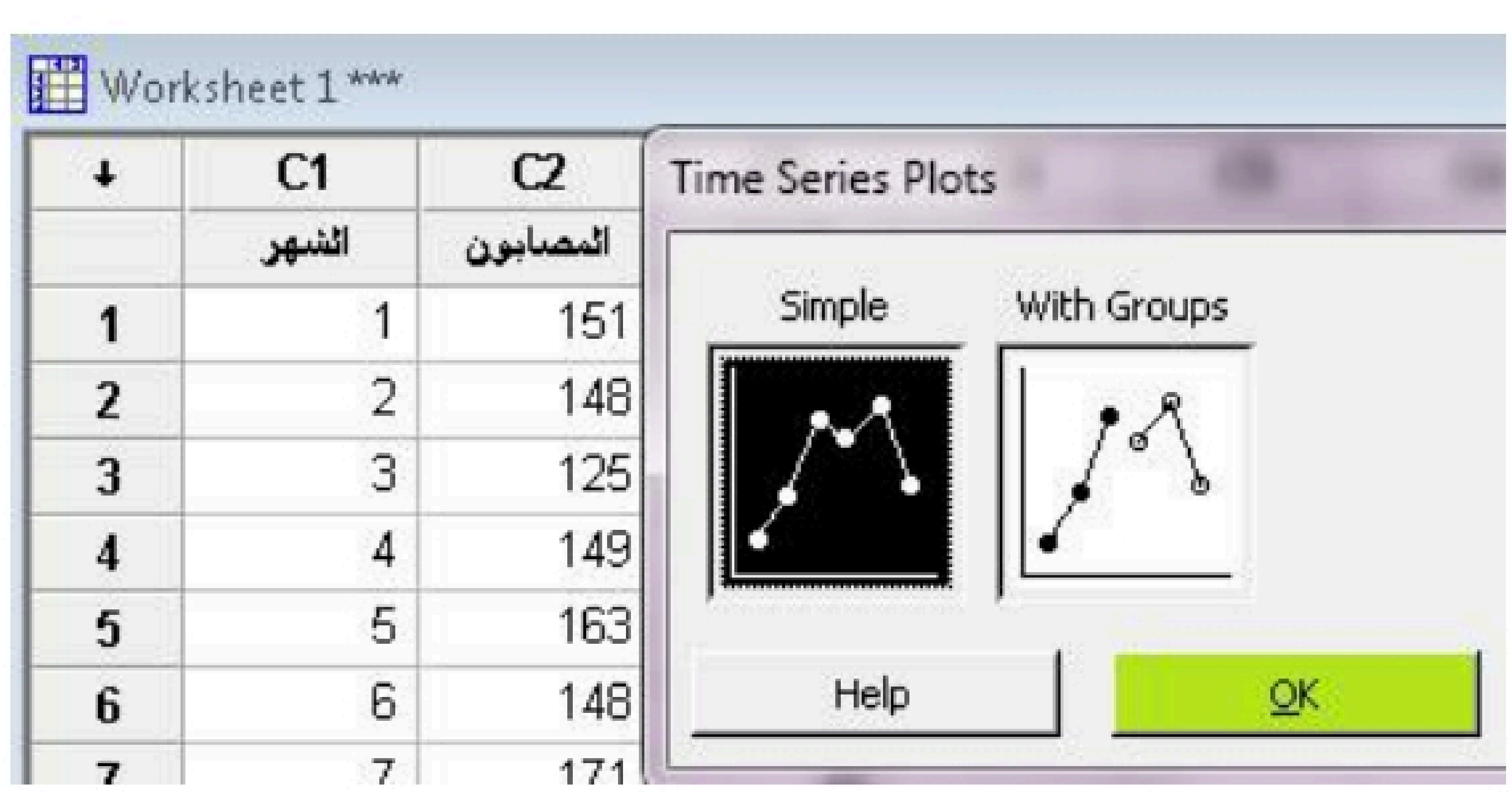

الشكل (١٢,١)

بالطبع من الممكن الحصول على بصمة المتسلسلة الزمنية باستخدام برنامج Minitab باستخدام الخطوات الآتية:

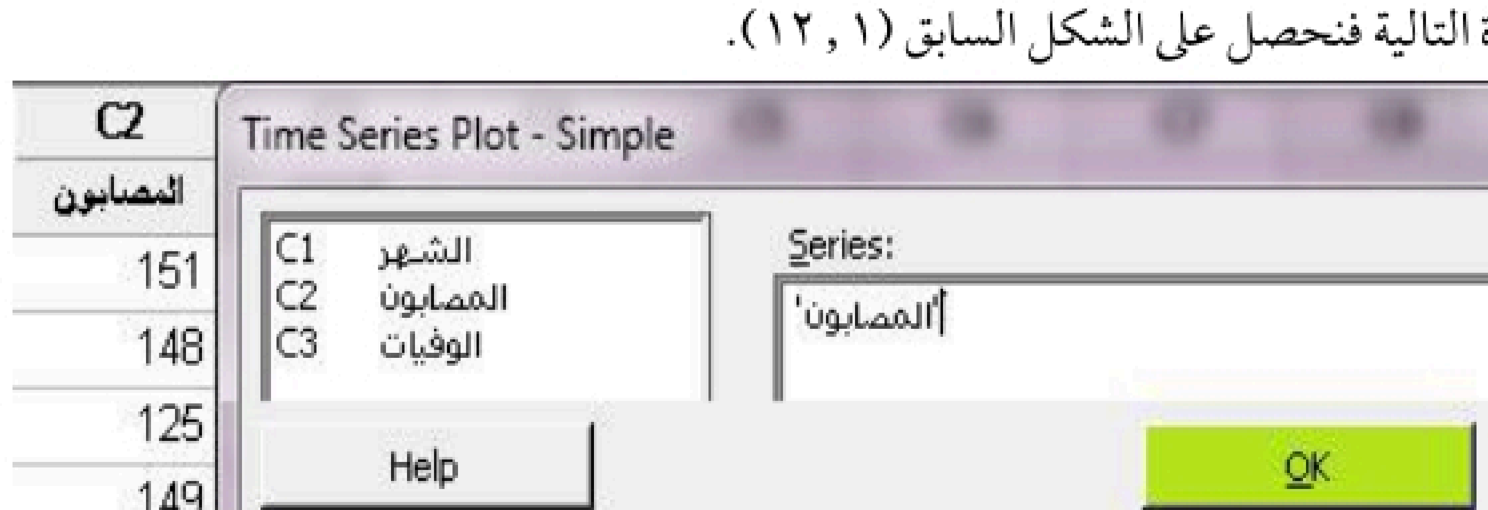
**Graph → Time Series → Simple → OK → Insert (المصابون) → OK**




ومن ثمَّ




وبعد ذلك نتَّبع الخطوة التالية فنحصل على الشكل السابق (١٢,١).

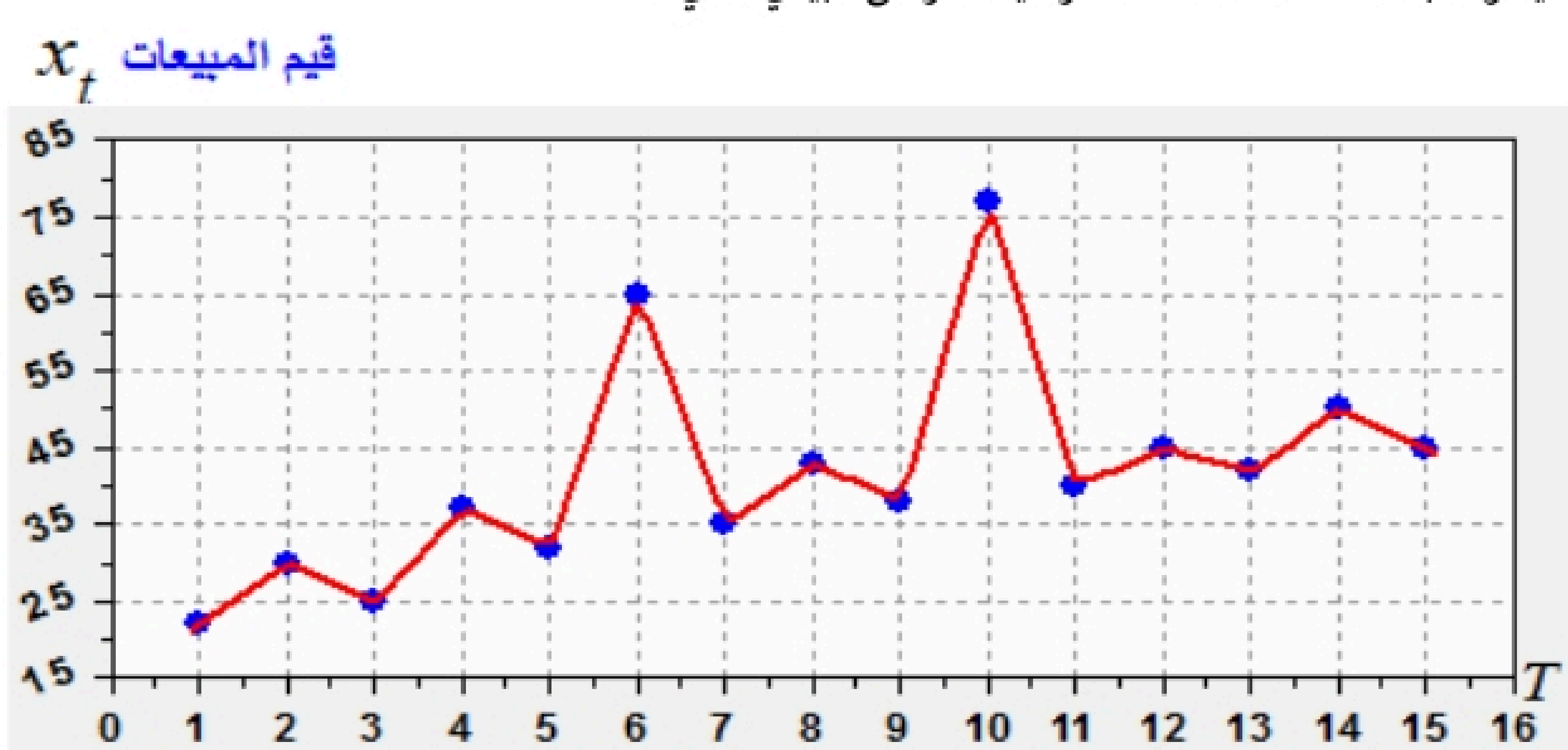


٢- لتكن لدينا المشاهدات الآتية التي تمثّل قيم المبيعات الشهرية (مقدّرة بـ 10000 وحدة نقدية) في متجر للملابس خلال خمسة عشر شهراً متتالياً.

الجدول (٢,١٢)

| الشهر $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|---|---|---|---|---|---|---|---|---|
| قيم المبيعات $x_t$ | 22 | 30 | 25 | 37 | 32 | 65 | 35 | 43 |
| الشهر $t$ | 9 | 10 | 11 | 12 | 13 | 14 | 15 | |
| قيم المبيعات $x_t$ | 38 | 77 | 40 | 45 | 42 | 50 | 45 | |

فعندئذ سيكون لبصمة هذه المتسلسلة الزمنية العرض البياني الآتي:



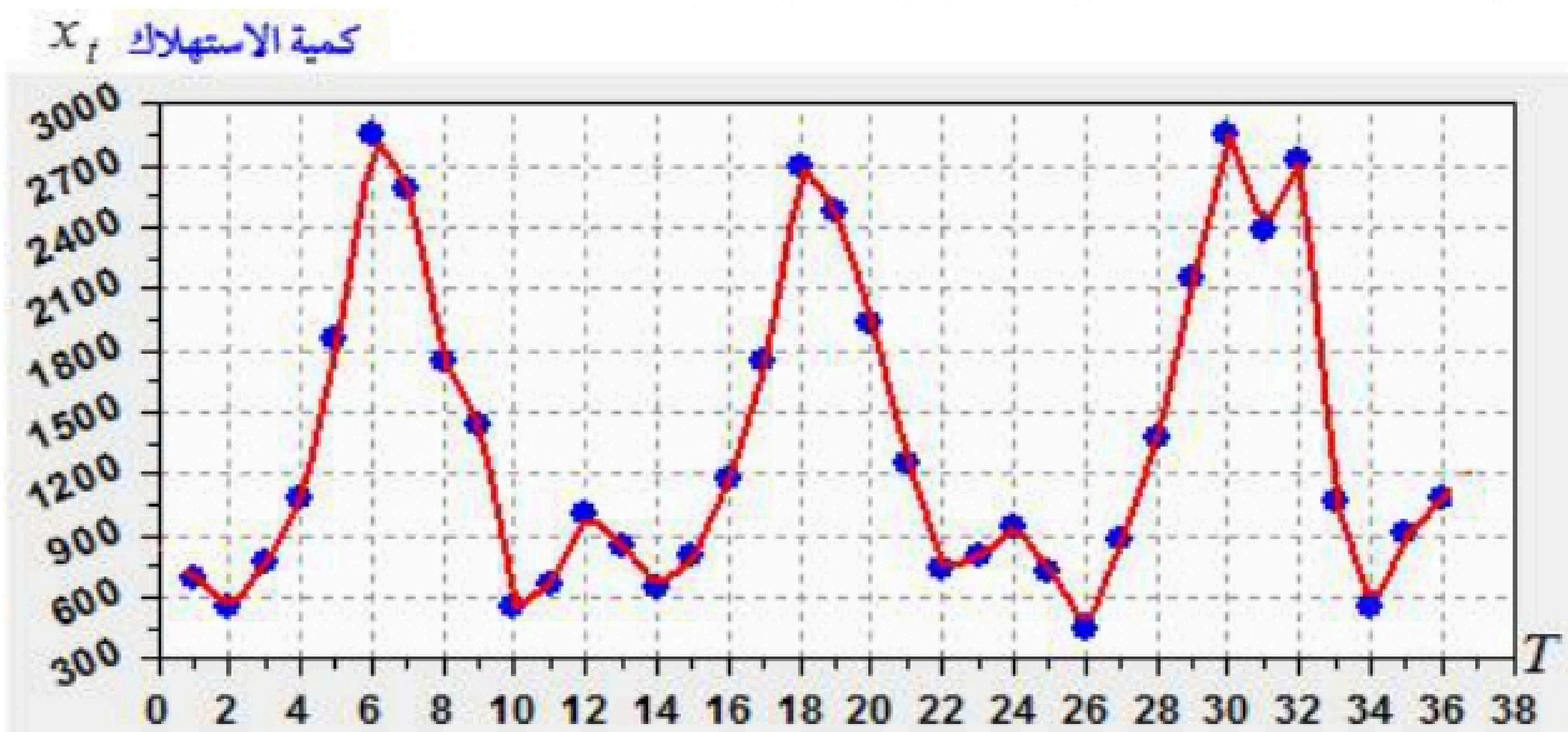

الشكل (٢,١٢)

٣- لتكن لدينا المشاهدات الآتية التي تمثّل كميّة الاستهلاك الشهري للكهرباء لإحدى الشقق السكنية خلال ثلاثة أعوام متتالية (مقدّرة ب الكيلو وات الساعي Kw/S).

الجدول (٣,١٢)

| تاريخ الفاتورة | كمية الاستهلاك | تاريخ الفاتورة | كمية الاستهلاك | تاريخ الفاتورة | كمية الاستهلاك |
|---|---|---|---|---|---|
| 1432/01 | 700 | 1433/01 | 850 | 1434/01 | 730 |
| 1432/02 | 555 | 1433/02 | 655 | 1434/02 | 455 |
| 1432/03 | 780 | 1433/03 | 805 | 1434/03 | 885 |
| 1432/04 | 1080 | 1433/04 | 1180 | 1434/04 | 1380 |
| 1432/05 | 1855 | 1433/05 | 1750 | 1434/05 | 2155 |
| 1432/06 | 2855 | 1433/06 | 2705 | 1434/06 | 2855 |
| 1432/07 | 2585 | 1433/07 | 2475 | 1434/07 | 2385 |
| 1432/08 | 1750 | 1433/08 | 1930 | 1434/08 | 2735 |
| 1432/09 | 1445 | 1433/09 | 1260 | 1434/09 | 1065 |
| 1432/10 | 560 | 1433/10 | 740 | 1434/10 | 550 |
| 1432/11 | 670 | 1433/11 | 800 | 1434/11 | 910 |
| 1432/12 | 1005 | 1433/12 | 950 | 1434/12 | 1085 |

فعندئذ سيكون لبصمة هذه المتسلسلة الزمنية العرض البياني الآتي:



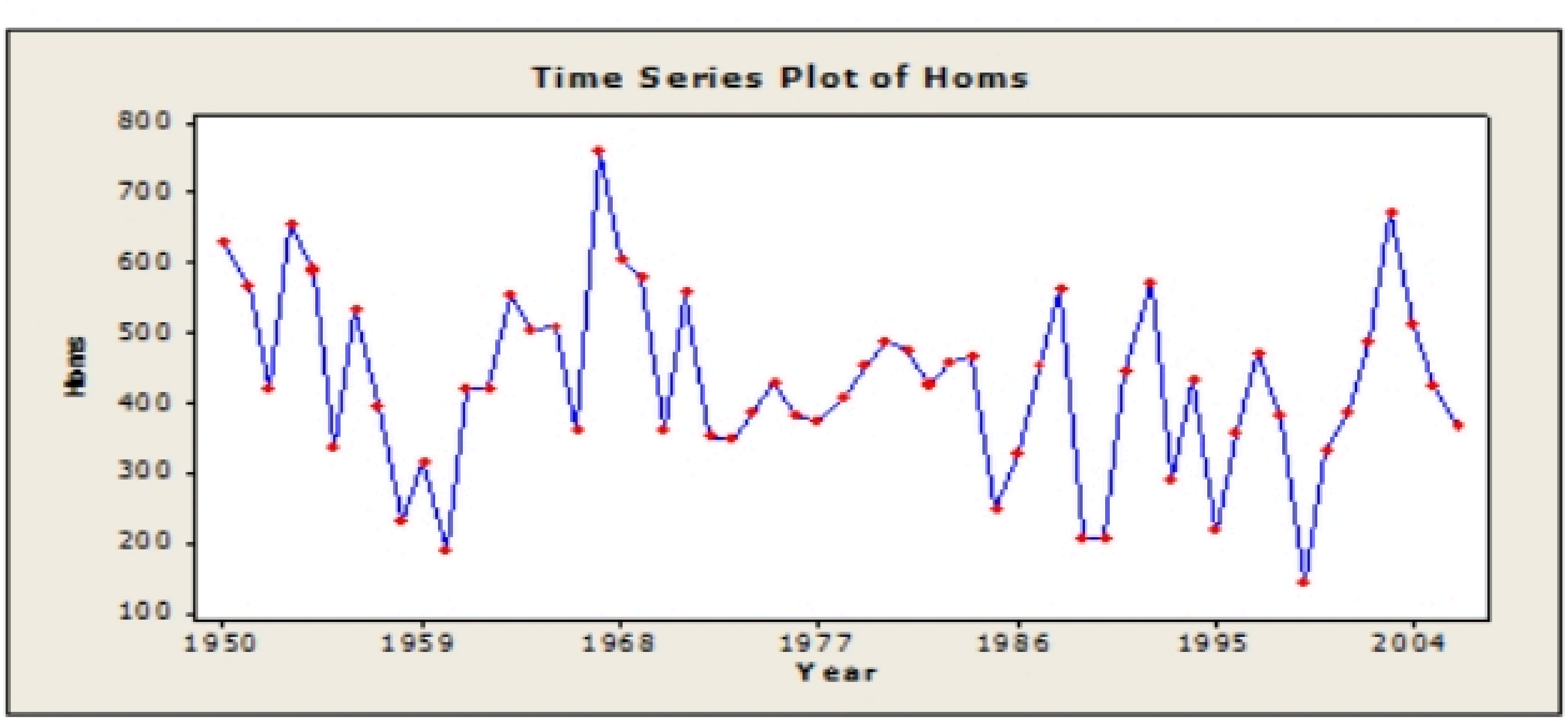

الشكل (١٢,٣)

٤- لتكن لدينا المشاهدات الآتية التي تمثّل متوسط كميّة الهطل المطري السنوي على مدينة (حمص في سوريا) في بلاد الشام خلال الفترة 1950–2004 (مقدّرة بالميليمتر على المتر المربّع $mm/m^2$).

الجدول (١٢,٤)

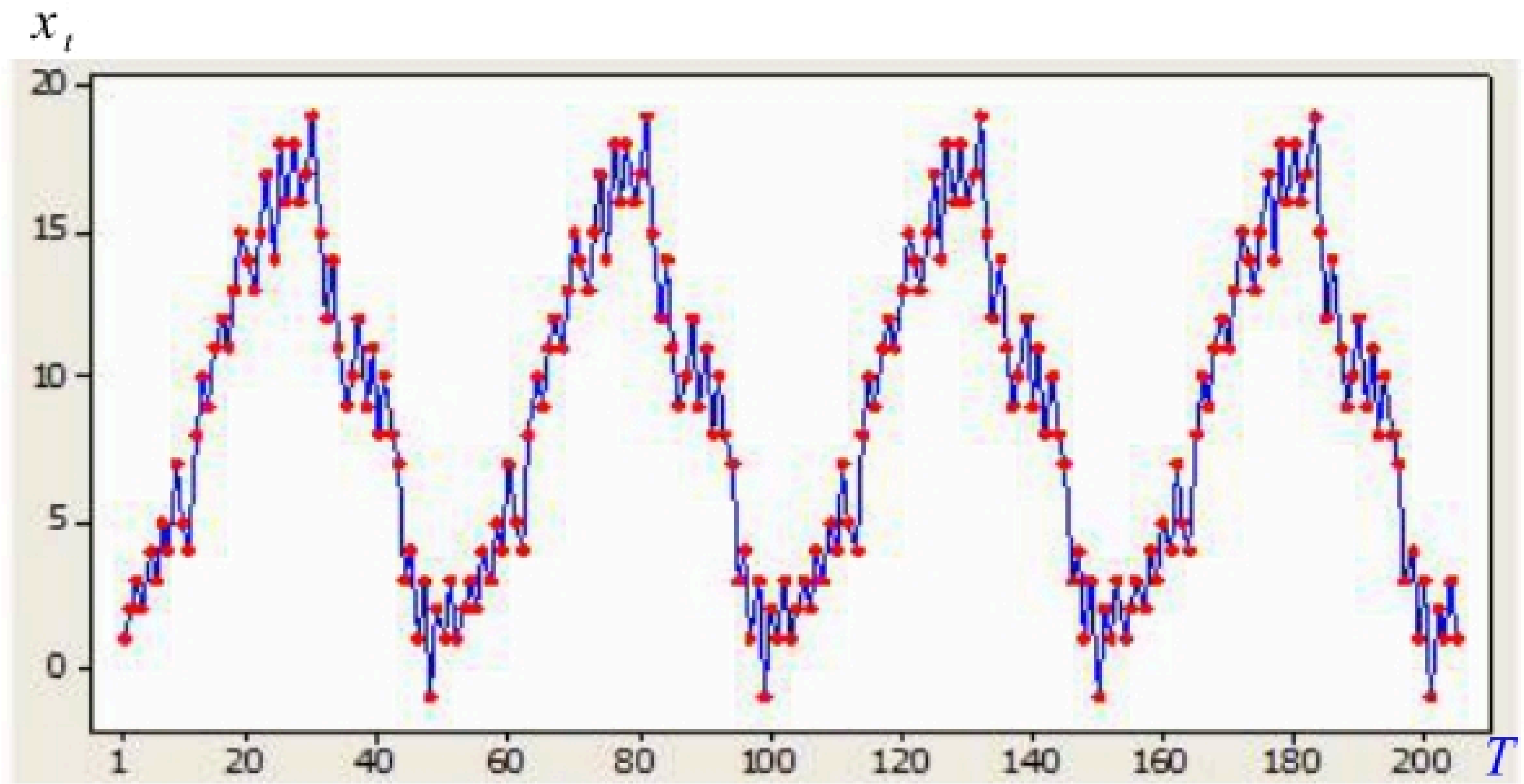
| كميّة الهطول المطري مقدّرة بالميليمتر على المتر المربّع $mm/m^2$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| السّنة | الكميّة | السّنة | الكميّة | السّنة | الكميّة | السّنة | الكميّة | السّنة | الكميّة |
| 1950 | 630 | 1961 | 420.5 | 1972 | 351.2 | 1983 | 455.3 | 1994 | 434 |
| 1951 | 566 | 1962 | 419.1 | 1973 | 346.2 | 1984 | 466.1 | 1995 | 216.7 |
| 1952 | 421.2 | 1963 | 552.9 | 1974 | 387 | 1985 | 249.1 | 1996 | 357.3 |
| 1953 | 654.5 | 1964 | 503.6 | 1975 | 429 | 1986 | 328.1 | 1997 | 468.4 |
| 1954 | 589.7 | 1965 | 506.4 | 1976 | 378.9 | 1987 | 450.3 | 1998 | 380.5 |
| 1955 | 333.1 | 1966 | 362.1 | 1977 | 375.2 | 1988 | 564.1 | 1999 | 143 |
| 1956 | 532 | 1967 | 758.9 | 1978 | 405.8 | 1989 | 205.5 | 2000 | 330 |
| 1957 | 395.2 | 1968 | 605.9 | 1979 | 452.2 | 1990 | 204.2 | 2001 | 388.1 |
| 1958 | 230.6 | 1969 | 578.6 | 1980 | 487.9 | 1991 | 446.9 | 2002 | 486.2 |
| 1959 | 314.8 | 1970 | 360.6 | 1981 | 474.6 | 1992 | 572 | 2003 | 672 |
| 1960 | 190 | 1971 | 556.5 | 1982 | 424.9 | 1993 | 289.7 | 2004 | 510.5 |

فعندئذ يكون لبصمة هذه المتسلسلة الزمنية العرض البياني الآتي:




الشكل (١٢,٤)

٥- لتكن لدينا المتسلسلة الزمنية الآتية (سنكتفي بعرض بصمة هذه المتسلسلة الزمنية بسبب كبر حجم العيِّنة $\tau = 205$ ، وهي متسلسلة زمنية افتراضية):




الشكل (٥ , ١٢)

الآن لو أمعنا النظر في العروض البيانية للمتسلسلات الزمنية السابقة، فإنَّنا سنلاحظ ما يلي:

١- لكلّ متسلسلة زمنية يوجد منحى يَرسم توجّهها على المدى البعيد للزمن. إنَّ هذا المنحى يُدعى **الاتجاه العام** Trend للمتسلسلة الزمنية، فعلى سبيل المثال نجد الاتجاه العام للمتسلسلتين الزمنيتين في المثالين /١/ و/٤/ هابطاً في حين أنَّ الاتجاه العام للمتسلسلة الزمنية في المثال /٢/ صاعداً، وأخيرا نلاحظ أنَّ الاتجاه العام للمتسلسلتين الزمنيتين في المثالين /٣/ و/٥/ يكاد يكون ثابتاً (أفقاً).

٢- يوجد في مسار معظم المتسلسلات الزمنية السابقة اهتزازات ذات أنماط مختلفة، فمنها:

أ- اهتزازات توافقية أو شبيهة بالاهتزازات التوافقية، كما يُلاحظ في الأشكال (١ , ١٢)، (٣ , ١٢) و (٥ , ١٢).

ب- اهتزازات متسعة المطال تحصل فجأةً وعند فترات مُحدّدة، كما يُلاحظ في الشكل (٢ , ١٢).

ج- اهتزازات صغيرة في مسار الاهتزازات السابقة، ولكنّها قد تكون غير منتظمة في تردّدها ومطالها، كما يُلاحظ في الشكل (٤ , ١٢) ، أو منتظمة في تردّدها ومطالها، كما يُلاحظ في الشكل (٥ , ١٢) .

إنَّ هذه الملاحظات توحي لنا أنَّ ثمَّة تركيبًا مُعيَّنًا لمكوِّنات ما أدَّت إلى هذا البناء للمتسلسلات الزمنية، بمعنى أنَّ بيانات المتسلسلات الزمنية هي في الواقع تفاعل لمجموعة مؤثِّرات يُمكنُ البحث في دراسة سلوك كلّ منها وفقاً للكائن الرياضياتي الذي يُمثِّلها. إنَّ هذا التأويل يدفعنا إلى معرفة البُنى الرياضياتية المكوِّنة لنموذج المتسلسلة الزمنية، وهذا ما ستقدِّمه لنا الفقرة الآتية.

### (٤ , ١ , ١٢) مُركِّبات المتسلسلات الزمنية Components of Time Series

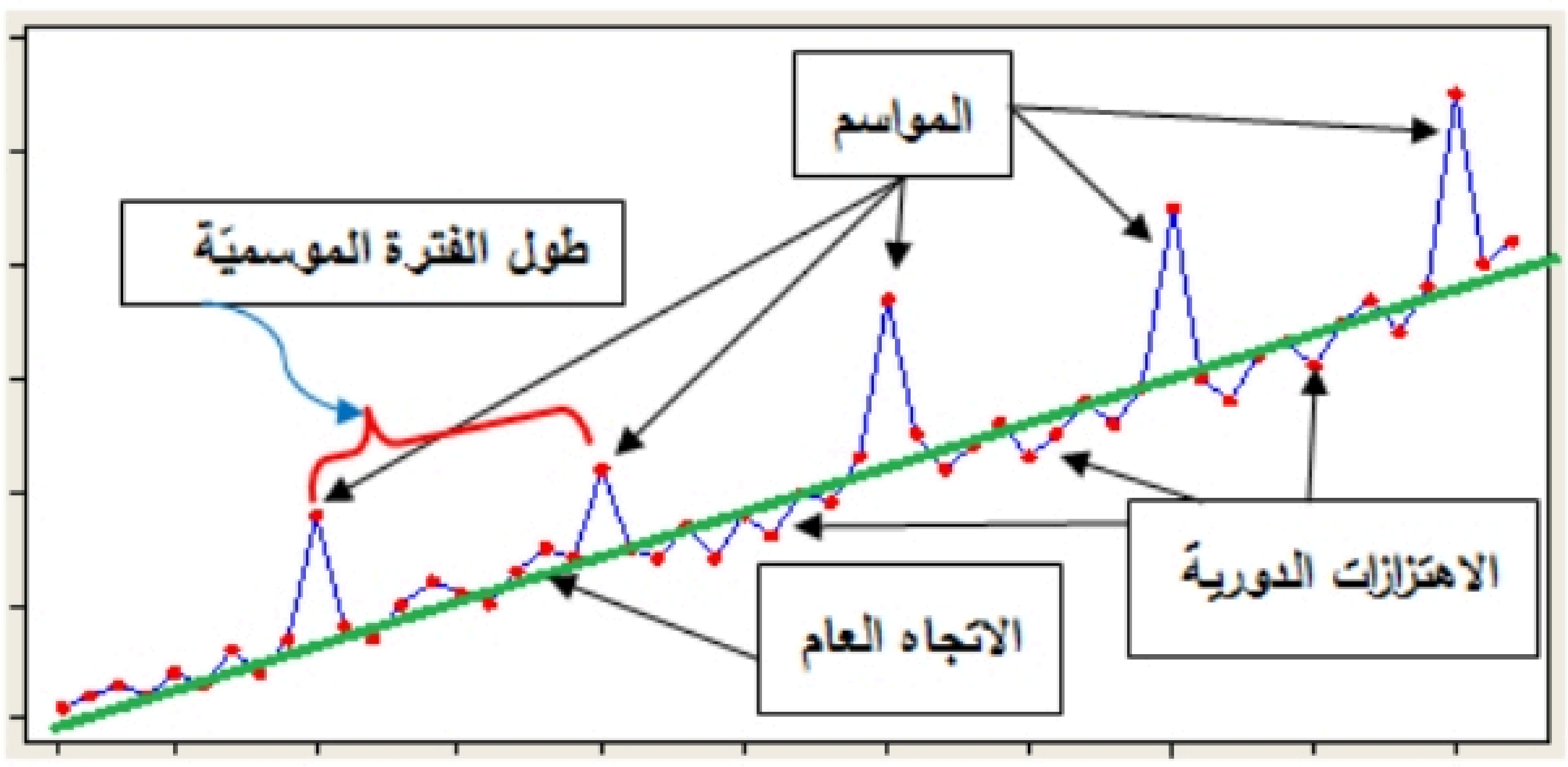
بناءً على ما سبق، فقد اقتُرِحَ تجزئة النموذج الرياضياتي المُمثِّل للمتسلسلة الزمنية إلى أربع مركبات هي:

١- **مُركِّبة الاتجاه العام** Trend Component، وسنرمز لها بـ $\mathbf{m}(t)$ وهي دالَّة بالزمن $t$، وهذه المركِّبة تصف التوجّه العام للمتسلسلة الزمنية على المدى البعيد للزمن.

٢- **المُركِّبة الموسميّة** Seasonal Component، وفي وصف هذه المركِّبة يوجد توجّهان:

أ- بعض المراجع (ومنها [33]) تصفها على أنَّها اهتزازات رتيبة Systematic Vibrations تحصل في مسار المتسلسلة الزمنية، ولكن هذه الاهتزازات تحصل في مواسم مُحدَّدة، كأن تظهر كل ساعة، أو كل خمس ساعات، أو كل شهر أو كل ربع سنة، أو...، وهذه

القيم الأخيرة (ساعة، خمس ساعات، شهر، ربع سنة أو...) تُدعى **طول الفترة الموسمِّية** (انظر الشكل الآتي)، ولكي نميِّزها عن التفسير الآتي في الفقرة / ب/ سندعوها **المركّبة الموسميَّة الصافية** ويُرمز لها بـ $\mathbf{s}(t)$، وهي دالَّة بالزمن $t$.




الشكل (٦ ,١٢)

ب- معظم البرامج الإحصائية وكذلك كثير من المراجع تصف المُركِّبة الموسميّة على أنَّها تلك المركّبة التي تشتمل على الخاصيّة السابقة / أ/ بالإضافة إلى التغيّرات (أو الاهتزازات) الرتيبة في مسار المتسلسلة الزمنية التي ليس لها انتظام موسمي، وهذه الاهتزازات الأخيرة تُمثّل في الواقع اهتزازات ذات نمط توافقي Harmonic Model شبيهة باهتزازات دالَّة الجيب **sin** وجيب التمام (التجب) **cos**، ولكي نميِّز هذه الاهتزازات عن التوصيف السابق / أ/ سنرمز لها بـ $\tilde{\mathbf{s}}(t)$، وهي دالَّة بالزمن $t$.

٣- **المُركِّبة الدوريّة** Period (or Cyclical) Component، وسنرمز لها بـ $\mathbf{p}(t)$ وهي دالَّة بالزمن $t$، وهذه المركّبة تصف التغيّرات (أو الاهتزازات) الرتيبة في مسار المتسلسلة الزمنية التي ليس لها نمط موسمي (كما ذكرناسابقاً)، وهذه المُركِّبة تمثّل الاهتزازات ذات الأنماط التوافقية أو الشبيهة بالتوافقية.

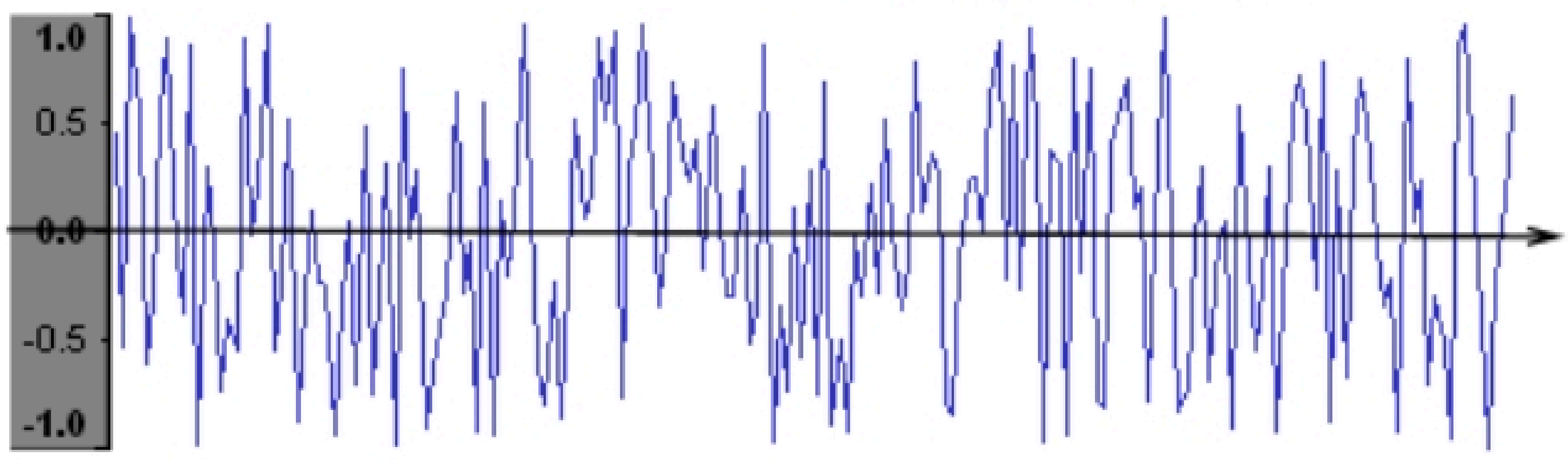
٤- **المُركِّبة العشوائية** (أو المُركِّبة غير المنتظمة) Irregular Component، وسنرمز لها بـ $N_t$، وهي متغيِّر عشوائي متعلّق بالزمن $t$، وتصف هذه المركّبة الاهتزازات غير الرتيبة في مسار المتسلسلة الزمنية، وقد تحتوي على مكوِّنات دوريّة غير منتظمة Irregular Cyclical ذات تردّد غير معلوم بالإضافة إلى اهتزازات أخرى لا يمكن تحديدها وفقاً للمركّبات السابقة، ومن ثَمَّ سيكون لهذه المركّبة سلوك عشوائي عند كل لحظة زمنية $t$، وبالتالي فإنَّ ممثّل هذه المركّبة سيكون طورياً عشوائياً. إنَّ الطوري العشوائي $\{N_t \,; t \geq 0\}$ الذي يمثّل الظاهر المذكورة آنفًا يكون عادة من نموذج **الضجيج الأبيض** White Noise، ولذلك يؤخذ المتغيِّر العشوائي $N_t$ على أنّه متغيِّر عشوائي بتوقع رياضياتي $\mathbf{E}N_t = 0$ وتباين يساوي $\mathbf{var}\,N_t = \sigma_e^2$ . علماً أنَّه يُقال عن طوري عشوائي $\{N_t\}_{t\geq 0}$ إنَّه طوري ضجيج أبيض White Noise Process إذا كان ما يلي مُحقَّقاً:

$$\mathbf{E}N_t = 0 \qquad ; t \geq 0 \tag{12,1-a}$$

$$\mathbf{cov}(N_s, N_t) = \begin{cases} \sigma_e^2 & for\ s = t \\ 0 & for\ s \neq t \end{cases} \qquad ; s,t \geq 0 \tag{12,1-b}$$

ويرمز لذلك بالشكل $N_t \sim \mathbf{WN}(0, \sigma_e)$، وعلاوة على ذلك يأخذ عادةً كمتغيّر عشوائي له توزيع طبيعي، وعندئذ يصبح للبند الثاني السابق المعنى الآتي:

من أجل كل القيم الممكنة لـ $t_1$ و $t_2$ مع $t_1 \neq t_2$ سيكون كل من المتغيّرين العشوائيين $N_{t_1}$ و $N_{t_2}$ مستقلان، والشكل الآتي يعرض لنا أحد مسارات هذا الطوري (طوري الضجيج الأبيض).




الشكل (١٢,٧)

لعلَّ المرء يتساءل في هذا الموضع عن سبب وجود المركَّبة العشوائية في نموذج المتسلسلة الزمنية، وكيف يُفسّر الافتراض أنَّ هذه المركَّبة من طبيعة الضجيج الأبيض؟

في الحقيقة إنَّ سبب وجود المركَّبة العشوائية في نموذج المتسلسلة الزمنية ناتج عن وجود أخطاء عديدة غير مُسيطر عليها تُرتكب أثناء أخذ القياسات المكوِّنة لبيانات المتسلسلة الزمنية، وهذه الأخطاء لا يمكن تفاديها أثناء عمليّة القياس مثل: الدقّة المحدودة لأجهزة القياس، أو الأخطاء الناتجة عن المجرّب، أو عوامل خارجية لا يمكن منعها في التأثير على الدراسة المعتمدة. أمّا تفسِّر الافتراض من أنَّ هذه المركَّبة هي من طبيعة الضجيج الأبيض (تشبيهاً بالضوء الأبيض) فإنَّ ذلك يعود إلى أنَّ هذه الأخطاء التي ذكرناها سابقاً قد تتداخل بعضها مع بعض وتولِّد أنماطاً جديدة من الأخطاء مَثلها في ذلك مَثل الضوْء الأبيض الناتج عن تداخل الألوان الرئيسية للضوء.

### (١٢,١,٤,١) ملاحظات

١- كما ذكرنا قبل قليل بأنَّ مراجع كثيرة وكذلك البرامج الإحصائية تتعامل مع الُمركَّبتين الدوريّة والموسميّة الموصوفتين سابقاً على أنَّها مركَّبة واحدة فقط، وتصف ها بالُمركَّبة الموسميّة. لكن يجب الانتباه هنا إلى أنَّه في حال غياب المركَّبة الُموسميّة الصافية من مسار المتسلسلة الزمنية فإنَّ المركَّبة الُموسميّة في هذه المراجع والبرامج الإحصائية هي في الواقع الُمركَّبة الدوريّة.

٢- في دراستنا لتحليل المتسلسلة الزمنية سنعتمد على التحليل الرباعي (أي وفقاً للمركبات الأربع) للمتسلسلة الزمنية وذلك بغية الزيادة في التوضيح لمركَّبات المتسلسلة الزمنية، وأمّا لدى استخدام برنامج Minitab فإنَّنا سنستخدم التحليل الثلاثي (أي وفقاً للمركبات الثلاث ) للمتسلسلة الزمنية فقط؛ لأنَّه يتعامل مع المركبتين الدوريّة والموسميّة على أنَّها مركَّبة واحدة.

٣- عندما نقوم بتوليد تنبؤات في الأمثلة القادمة سوف نقدِّم ونعرض هذه التنبؤات بدون إضافة المركَّبة العشوائية وذلك على سبيل التوضيح، علماً أنَّ معظم البرامج الإحصائية تقدِّم فترات ثقة للقيم المتنبأ بها، وسنقدِّم إحداها في أحد الأمثلة القادمة.

**الآن قد يتساءل القارئ لماذا نقوم بدراسة وتحليل المتسلسلات الزمنية؟**

إنَّ الهدف الأساس لدراسة وتحليل المتسلسلات الزمنية يكمن في التنبؤ عن قيم مستقبلية للبيانات المقدَّمة، وهذا يقودنا بدوره إلى وضع الاستراتيجيات اللازمة للتحكم (إذا أمكن ذلك) في المسار المستقبلي للمتسلسلة الزمنية بحيث لا يتجاوز حدود معيّنة، وذلك من خلال فرض شروط معينة على مَعَالم النموذج الممثّل للمتسلسلة الزمنية. إنَّ تحقيق ذلك يتطلب الحصول على وصف دقيق للملامح الخاصّة للطوري العشوائي الذي سيولّد المتسلسلة الزمنية، ومن ثمَّ إنشاء نموذج لتفسير وشرح سلوك المتسلسلة بدلالة متغيِّرات أخرى تربط القيم المشاهدة ببعض قواعد سلوك المتسلسلة الزمنية.

**إذا تمَّ توليد تنبؤات من النموذج المستنتج لمتسلسلة زمنية، فهل هذه التنبؤات المولَّدة ذات مصداقية؟**

في الحقيقة إنَّ سبب ثقتنا بقبول تنبؤات النموذج المستنتج يعود إلى أنَّ المتسلسلة الزمنية هي مجموعة مشاهدات تتولّد على التتالي مع مرور الزمن، وأنّ هذه المشاهدات المتتالية غالباً ما تكون مرتبطة (غير مستقلَّة إحصائياً) بحيث يعتمد بعضها على البعض الآخر، وهذا الارتباط العشوائي بين البيانات سيولِّد لدينا الثقة من أنَّ المتسلسلة الزمنية ستتابع مسيرتها في المستقبل على النحو الذي سلكته من الماضي وحتى الحاضر (وفقاً للدالَّة التي تمثِّل الاتجاه العام لها).

**هل يمكن استخدام النموذج الرياضياتي لمتسلسلة زمنية لتقدير قيم من الماضي؟**

في الواقع من الأغراض الأخرى لدراسة وتحليل المتسلسلات الزمنية هي معرفة سلوك البيانات أو التخمين (التقدير) لبعض القيم الواقعة ضمن نطاق البيانات، وذلك بشكل مشابه لما سبق أن ذكرنا في تحليل الانحدار، ولهذا السبب يلاحظ أنَّ تحليل الانحدار يُعدّ من المواضيع المهمّة لدراسة تحليل المتسلسلات الزمنية. بكلمات أخرى، فإنَّه يمكن استخدام النموذج الرياضياتي لمتسلسلة زمنية لتقدير قيم من الماضي على أن تكون واقعة ضمن نطاق البيانات المقدَّمة.

في الحقيقة إنَّ إيجاد نموذج مناسب لتمثيل متسلسلة زمنية مُعطاة يُعدّ من المهام الصعبة التي تحتاج إلى الكثير من البحث والتمحيص والخبرة. لقد ذكرنا أنَّ أول خطوة في دراسة المتسلسلة الزمنية هي تعيين بصمة هذه المتسلسلة، ومن ثمَّ تأتي الخطوة التالية المتمثِّلة باختيار نموذج رياضياتي مناسب للمتسلسلة الزمنية معتمدين في ذلك على بعض المقاييس الإحصائية التي تميّز نموذجًا عن آخر. لكن في الواقع توجد نماذج كثيرة للتمثيل الرياضياتي للمتسلسلات الزمنية، وهذه النماذج تتدرج ما بين البسيط نسبياً والمعقَّد، والغاية الأساس من جميع هذه النماذج هو توصيف الطوريات العشوائية الممثِّلة لهذه المتسلسلات الزمنية من أجل استخدامها في استنباط التنبؤات، والتحكّم في مسارها المستقبلي إن أمكن ذلك.

في كتابنا هذا سنقدّم، وبشيء من التفصيل، أحد النماذج البسيطة فقط؛ لكي يتبلور لدينا مفهوم المتسلسلة الزمنية بشكل مقبول، وسنترك التفاصيل المعقَّدة والنماذج التي تحتاج إلى مفاهيم متقدِّمة من نظرية الطوريات العشوائية مثل نماذج الانحدار الذاتي والمتوسطات المتحركة المتكاملة ARIMA وما يشتق عنها، ونماذج التنعيم Smoothing Models، و... أمّا الطريقة التي سنتناولها في دراسة تحليل متسلسلة زمنية فهي طريقة التجزئة فقط.

## (١٢,٢) طريقة التجزئة في تحليل متسلسلة زمنية

**Decomposition Method for the Analysis of Time Series**

إنَّ طريقة التجزئة (أو ما يُعرف باسم التجزئة الوصفية أيضًا) في تحليل المتسلسلات الزمنية تنطلق من خلال تجزئة العرض الخاص بالطوري العشوائي المولِّد للمتسلسلة الزمنية إلى عدّة مركبات، وبناءً على ذلك يمكن أن ينشأ لدينا نماذج عديدة. إنَّ النموذجين الأكثر استخداماً في تحليل المتسلسلات الزمنية وفقاً لطريقة التجزئة هما:

١ – نموذج الجمع Additive Model وتُستخدم من أجله العلاقة الآتية:

$$X_t = \mathbf{m}(t) + \mathbf{s}(t) + N_t \qquad [12,2\text{-a}]$$

وفي بعض المراجع يُعرض على النحو الآتي أيضاً:

$$X_t = \mathbf{m}(t) + \mathbf{p}(t) + \mathbf{s}(t) + N_t \qquad [12,2\text{-b}]$$

٢ – نموذج الضرب Multiplicative Model وتُستخدم من أجله العلاقة الآتية:

$$X_t = \mathbf{m}(t) \cdot \mathbf{s}(t) \cdot N_t \qquad [12,3\text{-a}]$$

وفي بعض المراجع يُعرض على النحو الآتي أيضاً:

[12,3-b] $$X_t = \mathbf{m}(t)\cdot\mathbf{p}(t)\cdot\mathbf{s}(t)\cdot N_t$$

علماً أنَّه لدينا في التمثيلات الأربع السابقة [12,2-a] وحتى [12,3-b] ما يلي:

$\mathbf{m}(t)$ هي دالَّة **الاتجاه العام** للمتسلسلة الزمنية،

$\mathbf{s}(t)$ هي الدالَّة التي تمثِّل **المركِّبة الموسميّة**، وفي حال التمثيل بأحد النموذجين [12,2-a] أو [12,3-a] فإنَّها تشتمل على الحركات الاهتزازية التي تحصل في مسار المتسلسلة الزمنية بالإضافة لبعض الاهتزازات التي تظهر في مواسم معيَّنة (المركِّبة الموسميّة الصافية)، ولكن دورها في التمثيلين [12,2-b] أو [12,3-b] يقتصر على المركِّبة الموسميّة الصافية التي تمثِّل التغيُّرات المتشابهة ذات التكرار المنتظم على مدى فترة من الزمن طولها هو **الدورة الموسميّة** Seasonal Cycle.

$\mathbf{p}(t)$ في التمثيلين [12,2-b] أو [12,3-b] هي الدالَّة التي تمثِّل **المركِّبة الدوريّة**، وهذه الدالَّة (كما ذكرنا سابقاً) تمثِّل الحركات الاهتزازية التي تحصل في مسار المتسلسلة الزمنية بدون المركِّبة الموسميّة الصافية.

$N_t$ هو متغيِّر عشوائي يُمثِّل **المركِّبة العشوائية** عند اللحظة الزمنية $t$، وبحيث يكون الطوري العشوائي $\{N_t\,; t\geq 0\}$ هو طوري ضجيج أبيض.

### (١٢,٢,١) ملاحظة

من الممكن ألا تحوي المتسلسلة الزمنية على المركبتين الدوريّة والموسميّة أو على إحداهما، بمعنى أنَّه من الممكن أن تكون المتسلسلة الزمنية ممثَّلةً باتجاهها العام مع الخطأ العشوائي فقط، وهذه الحالة تحصل عندما تكون بيانات المتسلسلة الزمنية قريبةً أو واقعةً على خط المنحنى الممثِّل للاتجاه العام للمتسلسلة الزمنية، والفقرتين /٢/ و /٣/ في المثال القادم (١٢,٢,١) توضِّحان لنا ذلك.

كما لاحظنا سابقاً فإنَّ دراسة المتسلسلات الزمنية تثير الكثير من التساؤلات، وهنا يظهر بعضاً منها أيضاً.

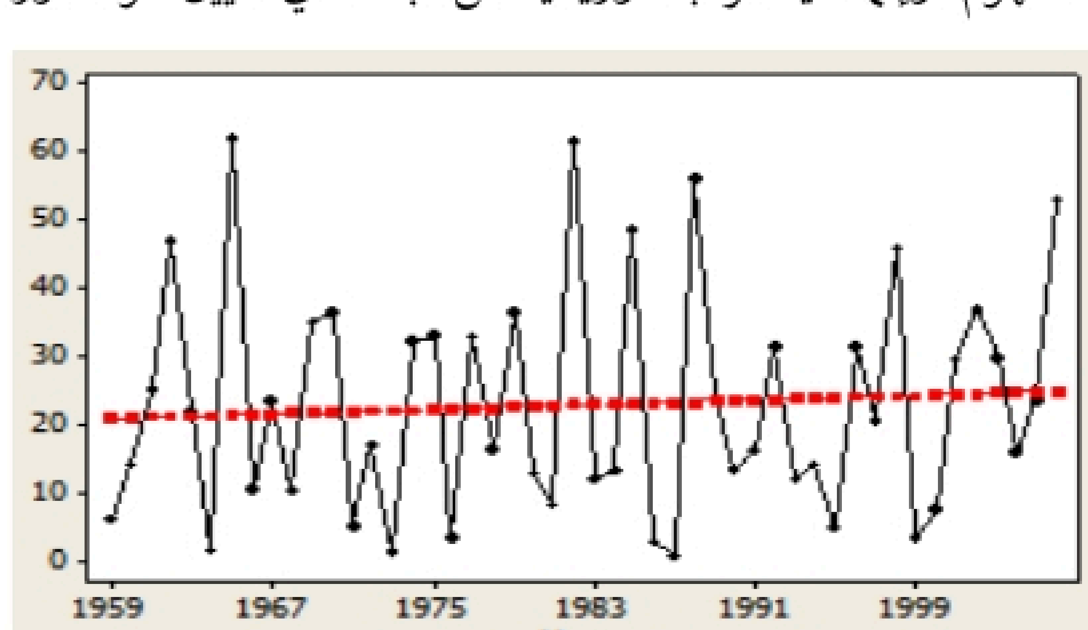
**متى نرغب باستخدام طريقة التجزئة؟**

في الحقيقة يمكننا استعمال طريقة التجزئة عندما نكون راغبين في تفريق المتسلسلة الزمنية إلى اتجاه عام خطي ومركِّبات موسميّة (دوريّة وموسميّة) بالإضافة إلى الخطأ، أو ببساطة إذا كنّا نريد أنْ نتفحص طبيعة الأجزاء المكوّنة للمتسلسلة الزمنية. حيث يمكننا اختيار نموذج الجمع، أو الضرب مع الاتجاه العام، وذلك حسب سلوك بيانات المتسلسلة الزمنية.

**متى نستخدم نموذج الجمع؟**

يُستعمل نموذج الجمع في طريقة التجزئة لتحليل المتسلسلات الزمنية عندما يكون الاتجاه العام غير مُحدّداً أو ثابتاً، وكذلك:

١- إذا كان للبيانات مطال (نمط) موسميٍّ ثابت (انظر العرض الآتي). حيث نلاحظ أنَّه في هذا العرض ليس لدينا مركِّبة موسميّة وفقاً للتعريف السابق لهذا المفهوم، وإنّما لدينا مركِّبة دوريّة يمكن البحث في تعيين طول دورها.




الشكل (١٢,٨)

٢- إذا كان حجم النمط الموسميّ غير متناسب مع البيانات (أو ليس نسبياً إلى البيانات)، بمعنى أنَّه في حال كان للبيانات اتجاه عام صاعد فإنَّه لا يترافق بزيادة مطّردة لمطالات المركّبة الموسميّة، وكذلك في حال كان للبيانات اتجاه عام هابط فإنَّه لا يترافق بنقصان مطّردة لمطالات المركّبة الموسميّة. بعبارة أخرى إنَّ مطالات المركّبة الموسميّة لا علاقة لها بالاتجاه العام للمتسلسلة الزمنية، فعلى سبيل المثال لو أمعنّا النظر في العروض الآتية فإنَّنا سنلاحظ في الشكل (٩, ١٢. أ) اتجاهًا عامًّا صاعدًا ومركّبة موسميّة شبه ثابتة (غير متأثرة الاتجاه العام) مع دورة موسميّة طولها يساوي 10 فجوات زمنية (10 Lags)، وكذلك في الشكل (٩, ١٢. ب) يُلاحظ وجود اتجاه عام هابط ومركّبة موسميّة شبه ثابتة بدورة موسميّة طولها يساوي 10 فجوات زمنية، ولدينا مركّبة دوريّة يمكن البحث عن طول دورها.


الشكل (٩, ١٢. أ)


الشكل (٩, ١٢. ب)


الشكل (٩, ١٢. ج)

لاحظ في هذا العرض الأخير (الشكل (٩ ,١٢ . ج)) أنّه لدينا اتجاه عام هابط ومركّبة موسميّة شبه ثابتة أيضاً مع دورة موسميّة بطول 10 فجوات زمنية أيضاً، وكذلك لدينا مركّبة دوريّة يمكن البحث عنه طول دورها.

٣- كذلك استخدام نموذج الجمع إذا كنّا نتطلع إلى الحصول على تنبؤات بعيدة المدى، ويُقصد بتنبؤات بعيدة المدى الحصول على تنبؤات لفترة زمنية قد تصل إلى طول دور كامل من أدوار المتسلسلة الزمنية.

**متى نستخدم نموذج الضرب؟**

يُستعمل نموذج الضرب في طريقة التجزئة لتحليل المتسلسلات الزمنية عندما نتطلع إلى تنبؤ بعيد المدى (كما هو الحال لدى النموذج الجمعي)، وكذلك إذا كان الاتجاه العام غير مُحدّد أو ثابت، أو إذا كان للبيانات نمط (شكل) موسميٍّ ذات مطال غير ثابت ومتأثر بالاتجاه العام، بمعنى أنَّ المطالات الموسميّة تتزايد في حالة الاتجاه العام الصاعد وتتناقص في حالة الاتجاه العام الهابط (انظر العرضين الآتيين).


الشكل (١٠ ,١٢ . أ)


الشكل (١٠ ,١٢ . ب)

من المعلوم أنّه لدى دراسة مسار متسلسلات زمنية معقّدة وطويلة الأمد يصبح عرض مركّباتها من خلال دوال بسيطة أمراً غير ممكن أو على الأقل صعبًا جداً، وفي مثل هذه الحالات يمكن تعيين مركّبات المتسلسلة الزمنية بشكل تقريبي من خلال أخذ فترة زمنية مَحدودة من مسار المتسلسلة الزمنية ومن ثمَّ البحث في دالَّة مناسبة لها على هذه الفترة، وبعد ذلك التعويض عن القيم المقيسة الأصل $x_t$ (التي تمثِّل مشاهدات المتسلسلة الزمنية الأصل) بالقيم المقدَّرة $\hat{x}_t$ النّاتجة عن الدالَّة المقدَّرة عند تعويض $t$ بالقيم الممكنة لها فنحصل بذلك على متسلسلة زمنية جديدة مشاهداتها $\hat{x}_t$، وهذه المشاهدات تستخدم في عملية التنبؤ.

**(١٢،٢،٢) ملاحظات**

١- في دراستنا المقبلة سنركِّز على استخدام نموذج الجمع فقط من أجل تحليل متسلسلة زمنية وذلك لسهولة التعامل مع هذا النموذج.

٢- إنَّ الطوري العشوائي $\{N_t\,;\,t \ge 0\}$ الذي سيرد في المركِّبة العشوائية للنماذج التي تُمَثِّل المتسلسلات الزمنية هو طوري ضجيج أبيض دوماً ما لم نشر إلى خلاف ذلك صراحة، أي سيكون لدينا $N_t \sim \mathbf{WN}(0,\sigma_e)$ .

لنقم الآن بتقديم دراسة مركِّبات المتسلسلة الزمنية مبتدئين ذلك بالاتجاه العام $\mathbf{m}(t)$ ، ثمَّ تقدير المركِّبة الدوريّة $\mathbf{p}(t)$ ، وبعد ذلك معالجة المركِّبة الموسميّة $\mathbf{s}(t)$ ، وأخيراً إضافة المتغيِّر العشوائي $N_t$ إلى النموذج.

## (١٢،٣) تحليل الاتجاه العام لمتسلسلة زمنية

## Trend Analysis of a Time Series

إنَّ عملية تقدير الاتجاه العام لمتسلسلة زمنية تعتمد وبقدر كبير على تحليل الانحدار (الخطي، وغير الخطي)، ولذلك يُفترض على المرء أن يكون لديه بعض المعارف المتعلِّقة بتقنيات الانحدار الخطي وغير الخطي قبل البدء بدراسة وتعيين الاتجاه العام لمتسلسلة زمنية.

لقد قمنا في الفصل الثالث بشرح الانحدار الخطي و معالجة بعض الحالات التي يمكن ردّها إلى خطيّة، وذكرنا أنَّ هناك حالات من الانحدار غير الخطي لا يمكن (أو من الصعب) خطّيتها، ومنها على سبيل المثال لا الحصر كثيرات الحدود من الدرجة $2 \le n$ .

سنقدِّم من خلال هذه الفقرة بعض الطرائق (البسيطة والعمليّة) المستخدمة في تعيين الدالَّة المقدِّرة للمسار (أو التوجّه) الذي تسلكه مجموعة بيانات عيِّنة مأخوذة من مسار متسلسلة زمنية مُعطاة، وسيكون هذا التقديم من خلال أمثلة تَرِد تباعاً.

**(١٢،٣،١) تقدير الاتجاه العام بوساطة مستقيم**

يمكننا في الواقع تقدير الاتجاه العام لمتسلسلة زمنية بوساطة مستقيم، وذلك باستخدام تحليل الانحدار الخطي البسيط كما قُدِّم في الفصل الثالث من هذا الكتاب. لذلك سنلاحظ أنَّ تحليل الانحدار سيلعب دوراً مهمًّا في تحليل الاتجاه العام لمتسلسلة زمنية.

**(١٢،٣،١،١) مثال**

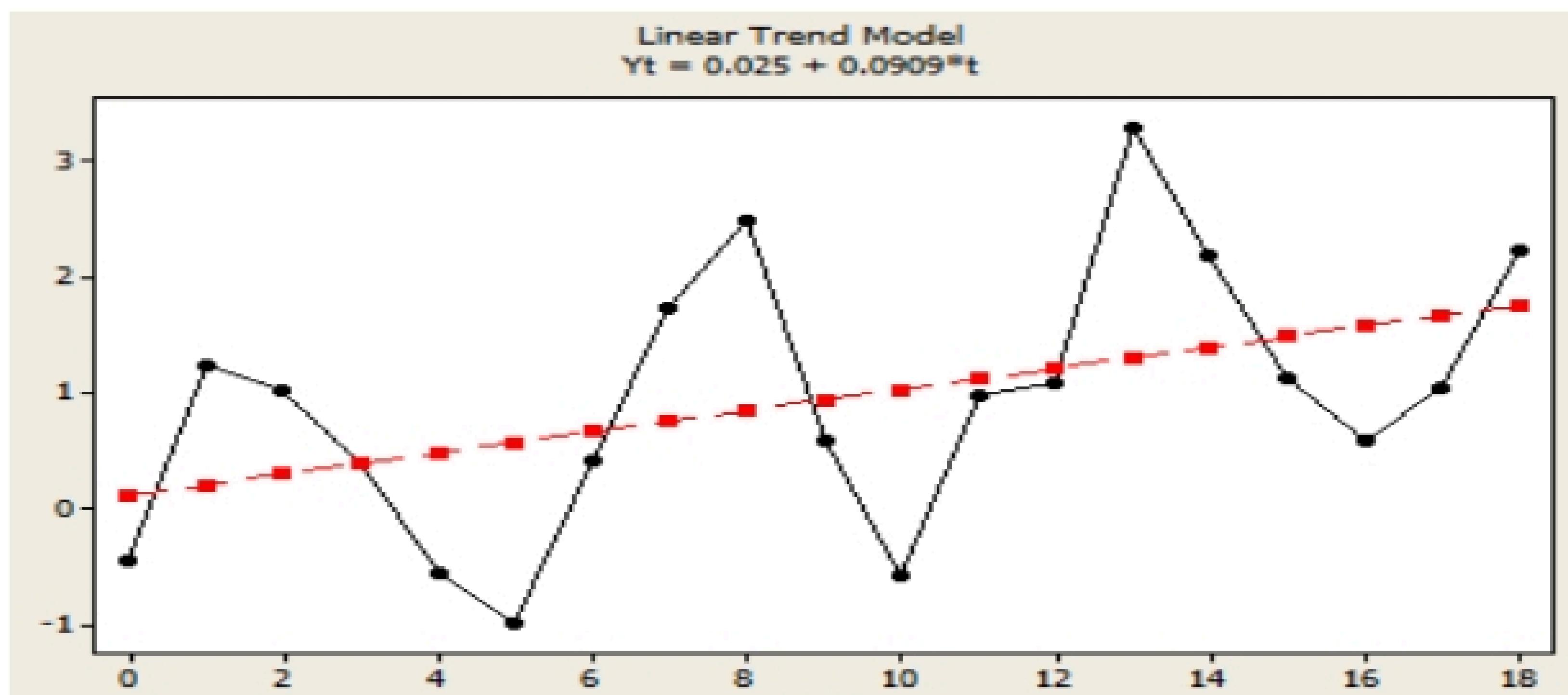
لتكن لدينا البيانات الآتية لعينة مأخوذة من مسار متسلسلة زمنية مولَّدة من طوري عشوائي $\{X_t\,;\,t \ge 0\}$ مُعطى بالعلاقة الآتية:

$$X_t = 0.1\,t + \sin(2\pi/6)t + N_t$$

الجدول (١٢،٥)

| Lag | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | -0.45 | 1.24 | 1.02 | 0.39 | -0.55 | -0.99 | 0.42 | 1.72 | 2.47 | 0.58 |
| Lag | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | |
| $x_t$ | -0.57 | 0.98 | 1.08 | 3.28 | 2.18 | 1.012 | 0.58 | 1.03 | 2.22 | |

فنجد أنَّ لبصمة بيانات هذه العيِّنة الشكل الآتي وموضَّحاً عليه الاتجاه العام الخطي:



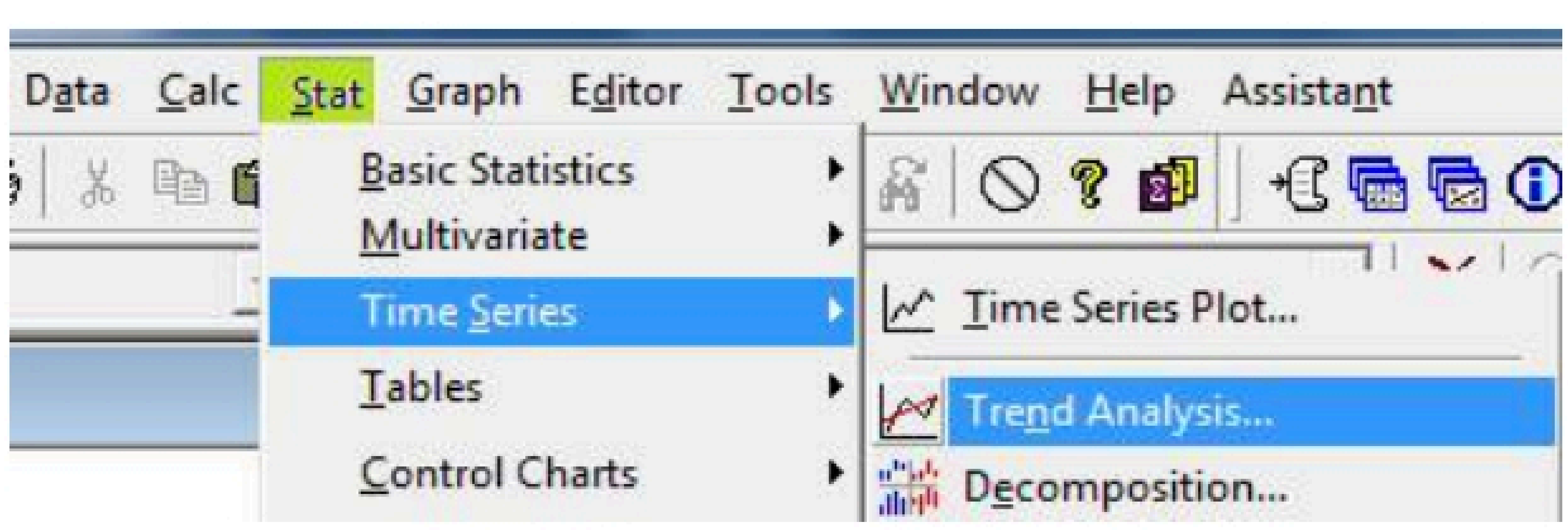

الشكل (١٠, ١٢)

علماً أنَّ معادلة مستقيم المربعات الصغرى الممثِّل للاتجاه العام الخطي لها العرض الآتي:

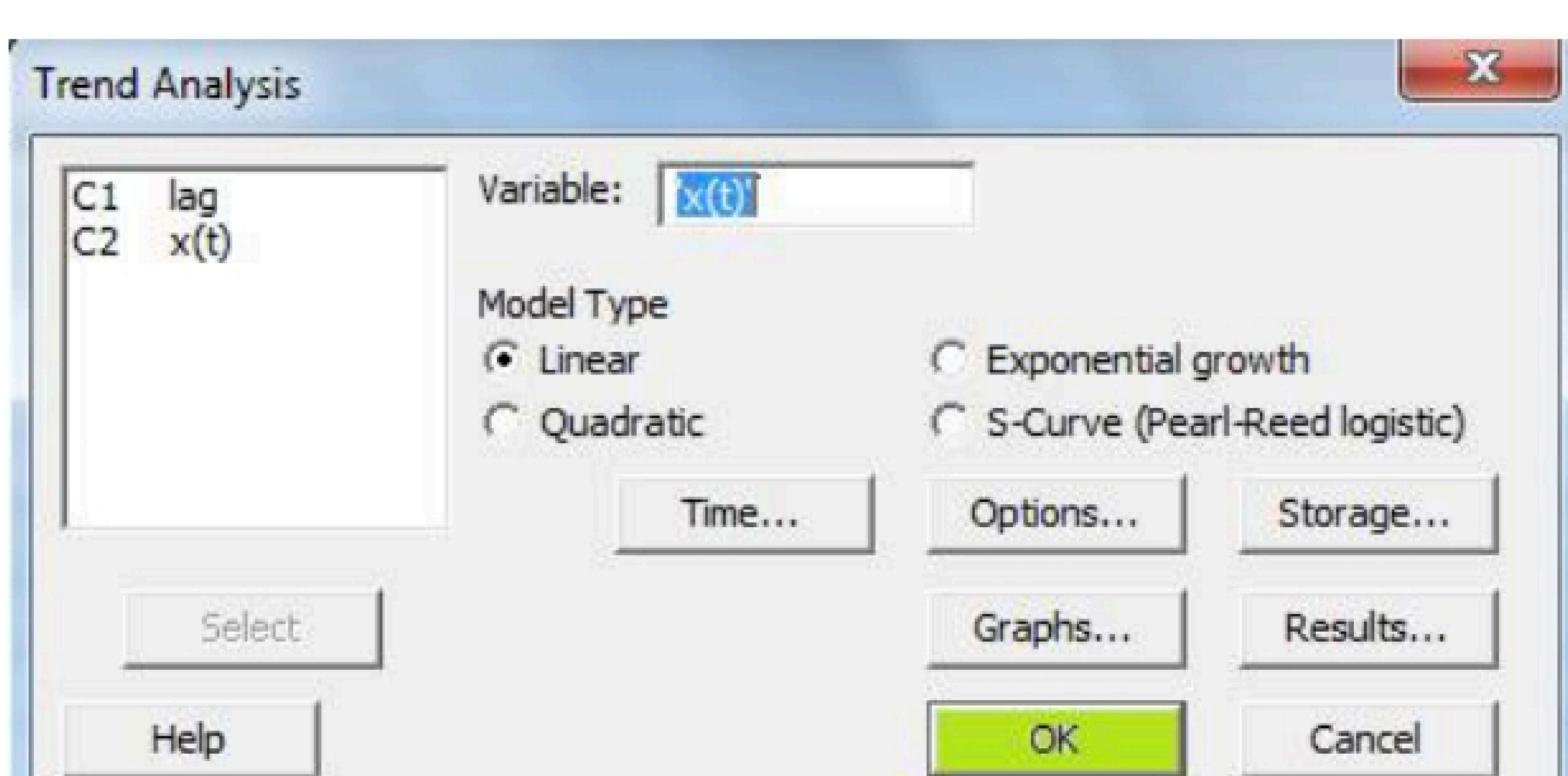
$$\mathbf{m}(t) = 0.025 + 0.0909\, t$$

ومن الممكن استخدام برنامج Minitab للحصول على الاتجاه العام الخطي السابق وذلك وفقاً للخطوات الآتية (والموضَّحة بالرسوم التالية أيضاً):

**Stat → Time Series → Trend Analysis → For Variable insert $x_t$ in free space → Activation Linear → OK.**




ومن ثَمَّ




ويمكننا هنا تغيير وحدة القياس وبعض الخصائص من خلال أيقونة Time و Options فنحصل الشكل (١٠, ١٢).

### (١٢,٣,٢) تقدير الاتجاه العام بوساطة كثيرة حدود

قد يصادفنا في كثير من الحالات أنَّ الاتجاه العام لمتسلسلة زمنية له منحى يمكن تمثيله بكثيرة حدود من درجة ما يبحث في تعيينها. في الحقيقية إنَّ استخدام كثيرات الحدود (إن أمكن ذلك) في تقدير الاتجاه العام لمتسلسلة زمنية يُعطينا ميزة جيدةً، وهذه الميزة تكمن في إمكانية رفع درجة كثيرة الحدود حتى الحصول على نموذج مناسب للاتجاه العام.

من الطرائق المستخدمة في تقدير الاتجاه العام لمتسلسلة زمنية بوساطة كثيرات الحدود ما يُعرف باسم **طريقة الفروق المتتالية.** إنَّ طريقة الفروق المتتالية لتعيين الدالَّة الممثِّلة لكثيرة حدود لبيانات مُعطاة تُعدّ من الطرائق الجيدة والعمليّة في تقدير كثيرات الحدود لمتسلسلة زمنية، وسنوجز شرحها فيما يلي.

#### (١٢,٣,٢,١) طريقة الفروق المتتالية لتعيين كثيرة الحدود Successive Differences Method to Determine a Polynomial

إنَّ **طريقة الفروق المتتالية** لتعيين كثيرة الحدود المُقدِّرة للاتجاه العام لمتسلسلة زمنية $x_1$، $x_2$،...و $x_n$ (بيانات مرتبة زمنياً) تقوم على الأسس الآتية:

١ – نفترض أنَّ كثيرة الحدود المُقدِّرة للاتجاه العام للمتسلسلة الزمنية لها العرض الآتي:

$$\mathbf{m}(t) = a_o + a_1\, t + a_2\, t^2 + \dots + a_n\, t^n$$

٢ – نقوم بحساب الفروق المتتالية:

$$\Delta^k x_t = \Delta^{k-1} x_{t+1} - \Delta^{k-1} x_t \quad ; k \in \mathbb{N}\,, t \in \{1\,, 2\,, \dots,\ \tau - k + 1\}$$

علماً أنَّ $k$ هو عدد الفروق المتتالية المستخدمة و $\tau$ عدد البيانات المُقدَّمة، وكذلك لدينا:

$$\Delta^1 x_t := \Delta x_t = x_{t+1} - x_t \qquad ; t = 1, 2, 3, \dots,\ \tau - 1$$

$$\Delta^0 x_t = x_t \quad ; t = 1, 2, 3, \dots,\ \tau$$

٣– حساب قيم المتوسط والتباين للفروقات المتتالية (التي عددها $k$) من خلال العلاقات الآتية:

$$\overline{\Delta^k x_t} = \frac{1}{\tau - k} \sum_{t=1}^{\tau - k} \Delta^k x_t$$

$$S_k^2 := \frac{1}{\tau - k} \sum_{t=1}^{\tau - k} \left(\Delta^k x_t - \overline{\Delta^k x_t}\right)^2$$

٤ – نقوم بحساب المقادير الآتية:

$$V_k := S_k^2 \cdot \binom{2k}{k}^{-1}$$

ومن ثمَّ حساب النسب $\frac{V_k}{V_{k+1}}$ من أجل كل قيم $k$ المستخدمة في حساب الفروق المتتالية.

٥ – نقوم بأخذ أول قيمة لـ $k$ يكون من أجلها $\frac{V_k}{V_{k+1}}$ أقرب ما يكون من الواحد كتقدير لقيمة $n$ وسنرمز لها ب $\hat{n}$ .

٦ – بعد معرفة درجة كثيرة الحدود $n$ نقوم بتُعيِّن أمثال كثيرة الحدود $a_0$، $a_1$،...و $a_{\hat{n}}$ على النحو الآتي:

أ – نقوم باستخدام التحويل الآتي على الدليل $t$ (تنفيذ عملية انسحاب لمحور الزمن) من خلال وضع:

$$r := \begin{cases} t-\bar{t} = \left(t - \dfrac{1+\tau}{2}\right) & for\ \tau\ odd \\ 2(t-\bar{t}) = 2\left(t - \dfrac{1+\tau}{2}\right) & for\ \tau\ even \end{cases} \quad (1)$$

ب- بناءً على قيمة $r$ التي نحصل عليها يكون لدينا:

$$\hat{a}_k = \frac{\overline{\Delta^k x_t}}{k} \quad ;for\ k \geq 1 \quad (2)$$

$$\hat{a}_o = \frac{1}{\tau}\left[\sum_r x_r - \hat{a}_2 \sum_r r^2\right] \quad (3)$$

ج- نعوِّض عن $\hat{a}_k$ بما يساويها من **(2)** و**(3)** في الصيغة الآتية:

$$\hat{\mathbf{m}}(r) = \sum_{k=0}^{\hat{n}} \hat{a}_k r^k \quad (4)$$

علماً أنَّ $\hat{n}$ هي درجة كثيرة الحدود المُقدَّرة سابقاً.

٧- في هذه الخطوة الأخيرة نعوض عن كل $r$ بما يساويها من **(1)** في الصيغة **(4)** فنحصل على كثيرة الحدود المطلوبة.

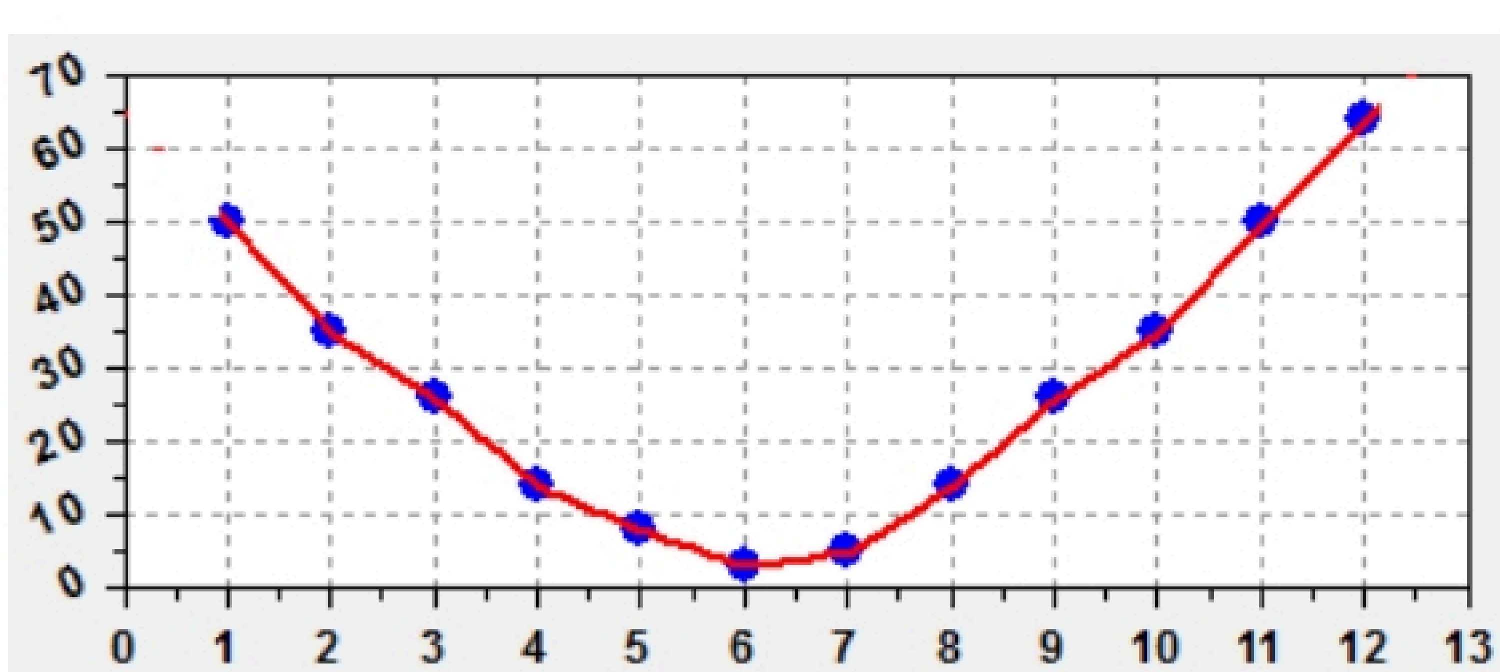
**(٢,٢,٣,١٢) مثال**

لتكن لدينا البيانات الآتية لعينة مأخوذة من مسار متسلسلة زمنية:

الجدول (٦,١٢.أ)

| Lag | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 50 | 35 | 26 | 14 | 8 | 3 | 5 | 14 | 26 | 35 | 50 | 64 |

فنجد أنَّ لبصمة هذه المتسلسلة الزمنية الشكل الآتي:




الشكل (١١,١٢.أ)

فنلاحظ في هذا العرض أنَّ المتسلسلة الزمنية المُعطاة لا تحتوي على مركَّبة موسميّة، وحتى المركَّبة الدوريّة فإنَّها ليست واضحة المعالم أيضاً، وكأنَّ الاتجاه العام هو المركَّبة غير العشوائية الوحيدة في هذه المتسلسلة الزمنية، وكذلك نلاحظ أنَّ مسارها له شكل بياني لدالَّة كثيرة حدود من درجة ما (**يبحث في تعيينها**).

الآن لتقدير الاتجاه العام لهذه المتسلسلة الزمنية سنستخدم طريقة الفروق المتتالية، ولذلك سنفترض أنَّ كثيرة الحدود المقدِّرة للاتجاه العام لهذه للمتسلسلة الزمنية هي:

$$\mathbf{m}(t)=a_o+a_1\,t+a_2\,t^2+\ldots+a_n\,t^n$$

ولنقم بحساب الفروقات المتتالية حتى المرتبة السادسة فنجدها كما في الجدول الآتي:

الجدول (٦, ١٢. ب)

| $s$ | -11 | -9 | -7 | -5 | -3 | -1 | 1 | 3 | 5 | 7 | 9 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| $x_t$ | 50 | 35 | 26 | 14 | 8 | 3 | 5 | 14 | 26 | 35 | 50 | 64 |
| $\Delta^1 x_t$ | -15 | -9 | -12 | -6 | -5 | 2 | 9 | 12 | 9 | 15 | 14 | |
| $\Delta^2 x_t$ | | 6 | -3 | 6 | 1 | 7 | 7 | 3 | -3 | 6 | -1 | |
| $\Delta^3 x_t$ | | -9 | 9 | -5 | 6 | 0 | -4 | -6 | 9 | -7 | | |
| $\Delta^4 x_t$ | | | 18 | -14 | 11 | -6 | -4 | -2 | 15 | -16 | | |
| $\Delta^5 x_t$ | | | -32 | 25 | -17 | 2 | 2 | 17 | -31 | | | |
| $\Delta^6 x_t$ | | | | 57 | -42 | 19 | 0 | 15 | -48 | | | |

ولنقم بحساب المقادير $V_k:=S_k^2\cdot\binom{2k}{k}^{-1}$، ومن ثَمَّ حساب النسب $\frac{V_k}{V_{k+1}}$ علماً أنَّ $k$ هو عدد الفروق المتتالية التي استخدمت، وقيم التباينات $S_k^2$ يجب أن تُحسب لكل من البيانات $x_t$ والفروقات $\Delta^k x_t$ المستخدمة في الجدول السابق، فنجدها من أجل مثالنا هذا كما في الجدول الآتي (٦, ١٢. ج):

الجدول (٦, ١٢. ج)

| قيم | المتوسط لقيم | التباين لقيم |
|---|---|---|
| $t$ | 6. 5 | |
| $x_t$ | 27.5 | 359.42 |
| $\Delta^1 x_t$ | 1.27 | 111.29 |
| $\Delta^2 x_t$ | 2.9 | 15.09 |
| $\Delta^3 x_t$ | -0.78 | 44.39 |
| $\Delta^4 x_t$ | 0.25 | 147.19 |
| $\Delta^5 x_t$ | -4.86 | 432.98 |

ومن ثَمَّ يكون لدينا ما يلي:

الجدول (٦, ١٢. د)

| $k$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $\binom{2k}{k}$ | 1 | 2 | 6 | 20 | 70 | 252 | 924 |
| $S_k^2$ | 359.42 | 111.29 | 15.09 | 44.39 | 147.19 | 432.98 | 1317.139 |
| $V_k$ | 359.42 | 55.645 | 2.515 | 2.220 | 2.102 | 1.817 | 1.425 |
| $\frac{V_k}{V_{k+1}}$ | 6.459 | 22.125 | 1.133 | 1.056 | 1.157 | 1.275 | |

وهكذا نجد أنَّ أول قيمة لـ $k$ يكون من أجلها $\frac{V_k}{V_{k+1}}$ أقرب ما يكون من الواحد هي $k = 3$، ومن ثمَّ تكون القيمة المُقدَّرة لـ $n$ هي $\hat{n} = 3$.

الآن وبعد معرفة درجة كثيرة الحدود التي يجب أخذها نقوم بتعيين مَعَالِم كثيرة الحدود $\mathbf{m}(t) = a_o + a_1 t + a_2 t^2 + a_3 t^3$ على النحو الآتي:

أ- نقوم بتحويل الدليل $t$ إلى دليل آخر $r$ من خلال وضع التحويل $r = 2(t - \bar{t})$ لأنَّ $\tau$ زوجي، فيكون لدينا ما يلي:

الجدول (٦, ١٢. هـ)

| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | sum |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $r$ | -11 | -9 | -7 | -5 | -3 | -1 | 1 | 3 | 5 | 7 | 9 | 11 | |
| $r^2$ | 121 | 81 | 49 | 25 | 9 | 1 | 1 | 9 | 25 | 49 | 81 | 121 | 572 |

ب- بناءً على قيمة $r$ المحسوبة في الجدول السابق يكون لدينا:

$$\hat{a}_1 = \overline{\Delta^1 x_t} = 1.27 \quad \& \quad \hat{a}_2 = \frac{\overline{\Delta^2 x_t}}{2} = \frac{2.9}{2} = 1.45 \quad \& \quad \hat{a}_3 = \frac{\overline{\Delta^3 x_t}}{3} = \frac{-0.78}{3} = -0.26$$

ومن ثمَّ ينتج لدينا:

$$\hat{a}_o = \frac{1}{\tau}\left[\sum_r x_r - \hat{a}_2 \sum_r r^2\right] = \frac{1}{12}[330 - 1.45 \times 572] = -41.62$$

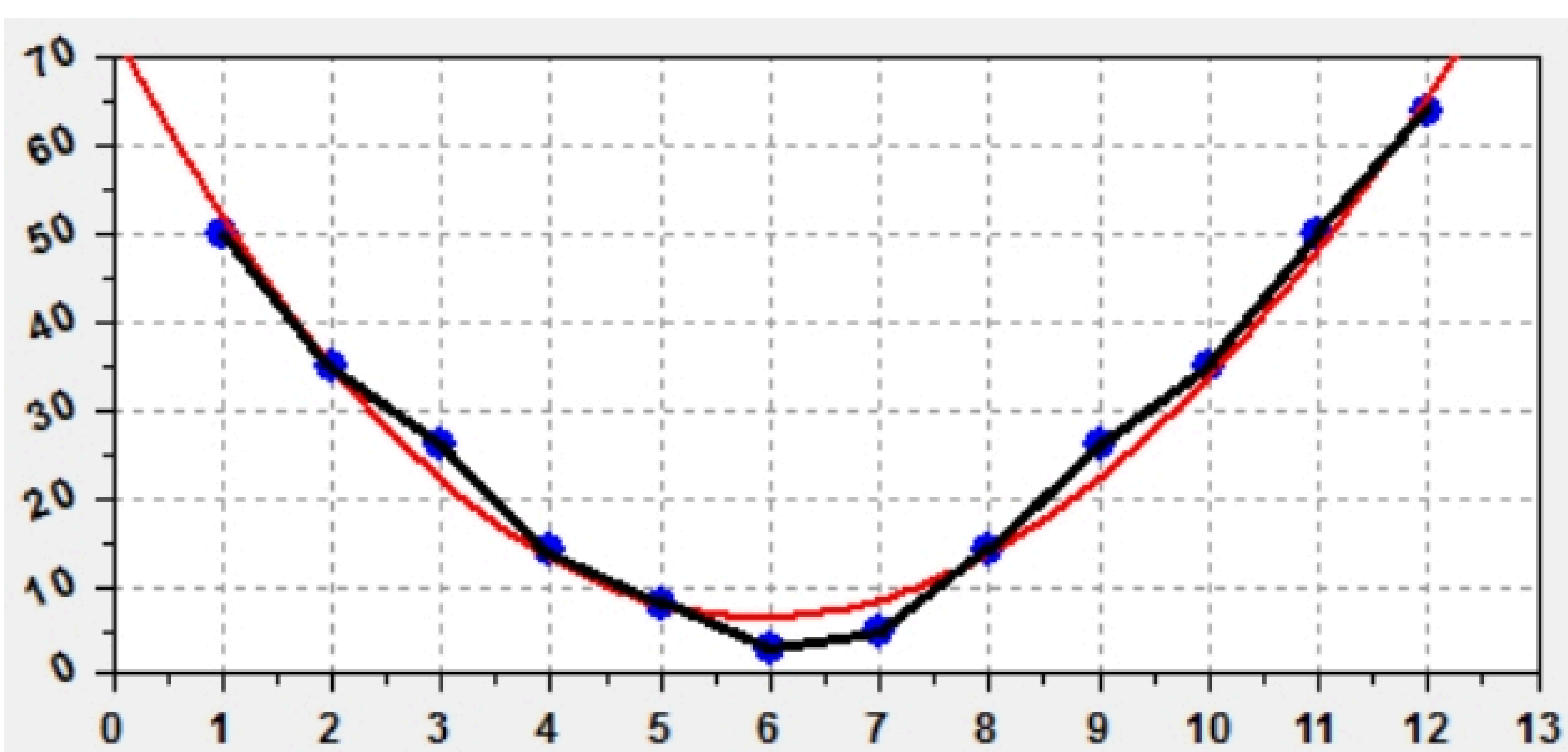
ج- نعوِّض في صيغة كثيرة الحدود $\hat{\mathbf{m}}(r) = \sum_{k=0}^{\hat{n}} \hat{a}_k r^k$ علماً أنَّ $\hat{n}$ هي درجة كثيرة الحدود المُقدَّرة سابقاً فنحصل على العرض الآتي:

$$\begin{aligned}\hat{\mathbf{m}}(r) &= -41.62 + 1.27\, r + 1.45\, r^2 - 0.26\, r^3 \\ &= -41.62 + 1.27\,[2(t-6.5)] + 1.45\,[2(t-6.5)]^2 - 0.26\,[2(t-6.5)]^3\end{aligned}$$

والتي نحصل منها أنَّ لكثيرة الحدود المقدَّرة للاتجاه العام لهذه المتسلسلة الزمنية العرض الآتي:

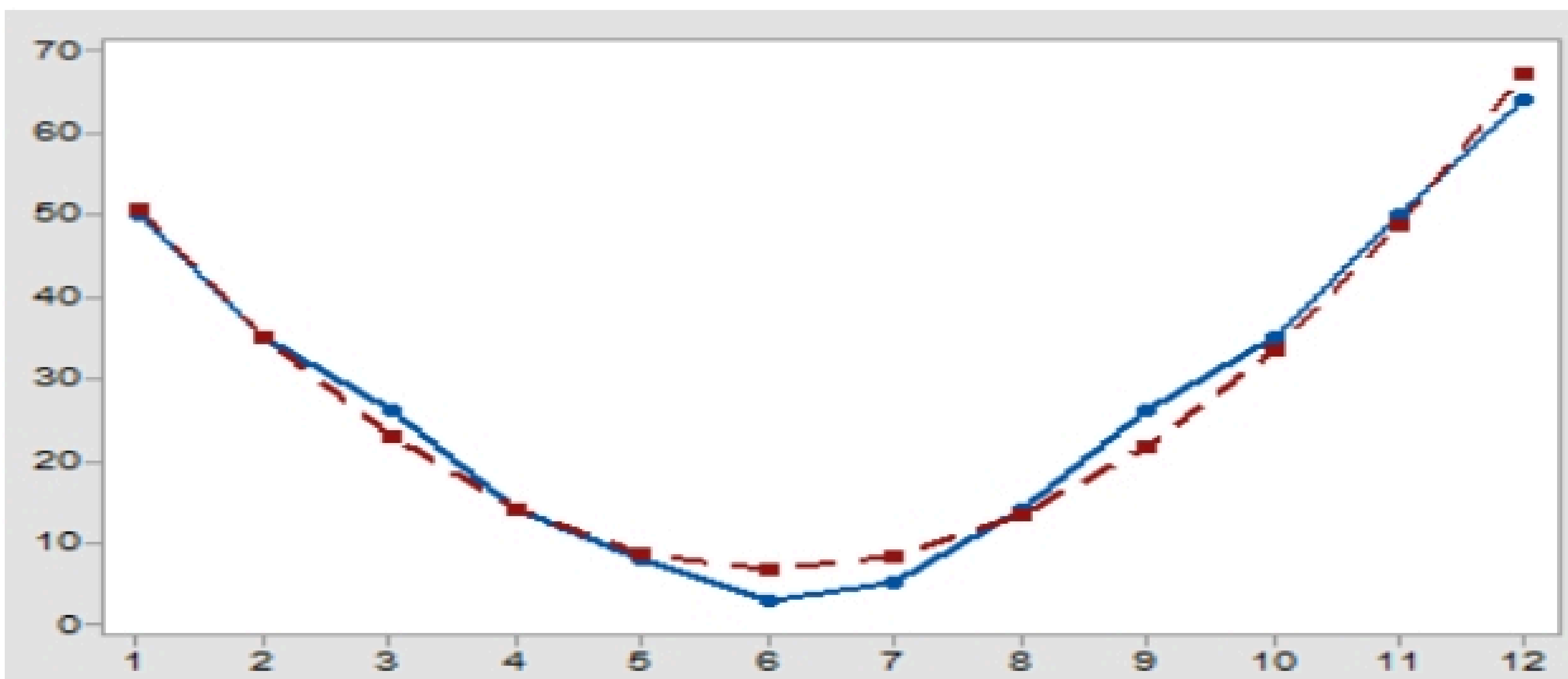
$$\hat{\mathbf{m}}(t) = 758.15 + 190.78\, t + 46.36\, t^2 + 2.08\, t^3$$

ولرسمها البياني الشكل الآتي.




الشكل (١١, ١٢. ب)

أمّا إذا أردنا استخدام برنامج Minitab في تقدير الاتجاه العام لهذه المتسلسلة الزمنية فإنَّه سيقدّم لنا كثيرة حدود من الدرجة الثانية فقط كتقدير مناسب للمشاهدات المُعطاة، ومعادلتها $\hat{\mathbf{m}}(t) = 69.64 - 20.76\,t + 1.7128\,t^2$، ولها العرض البياني الآتي:



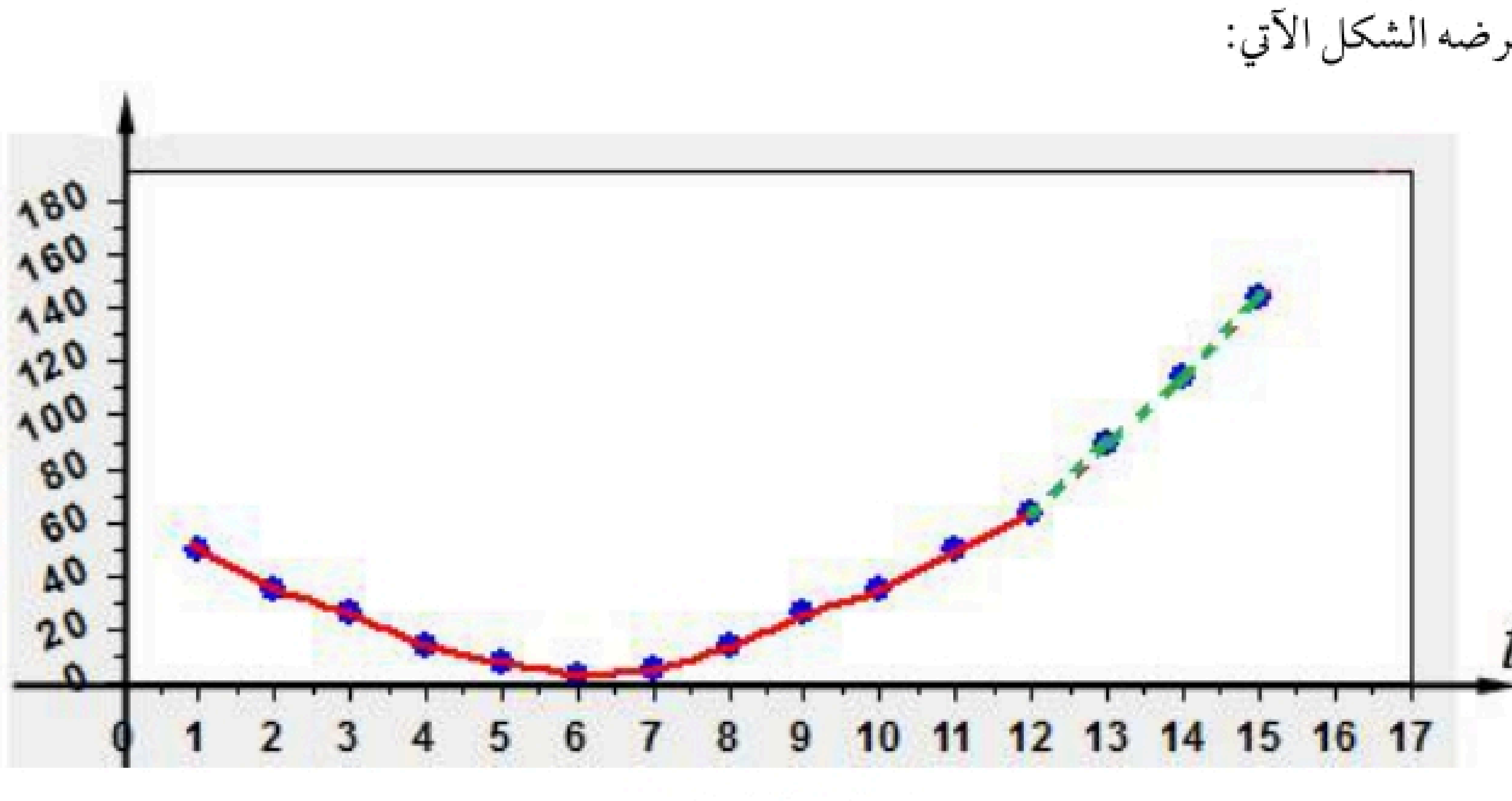

الشكل (١١ ,١٢.ج)

لاحظ جودة طريقة الفروق المتتالية في عملية التقدير بسبب إعطاء كثيرة حدود من درجة أعلى مما يقدّمه برنامج Minitab الذي أدى إلى توفيق أفضل للبيانات، ولكن هذا لا يُعدّ قاعدة عامّة، فلو كانت كثيرة الحدود الملائمة للبيانات هي كثيرة حدود من الدرجة الثانية فعلاً لوجدنا أنَّ النموذج الذي يقدّمه برنامج Minitab أكثر دقَّة من ذلك الذي تقدِّمه طريقة الفروق المتتالية. كما نلاحظ هنا أنَّ النقاط الممثِّلة لبيانات المتسلسلة الزمنية قريبة من منحنى الاتجاه العام، ولذلك يُمكننا وبخطأ صغير نسبياً استخدام معادلة منحنى الاتجاه العام $\hat{\mathbf{m}}(t)$ لهذه المتسلسلة الزمنية من أجل الحصول على تنبؤات جيدة، فمن أجل المتسلسلة الزمنية السابقة نحصل على التنبؤات المقدَّمة في الجدول الآتي:

الجدول (٦ ,١٢.و)

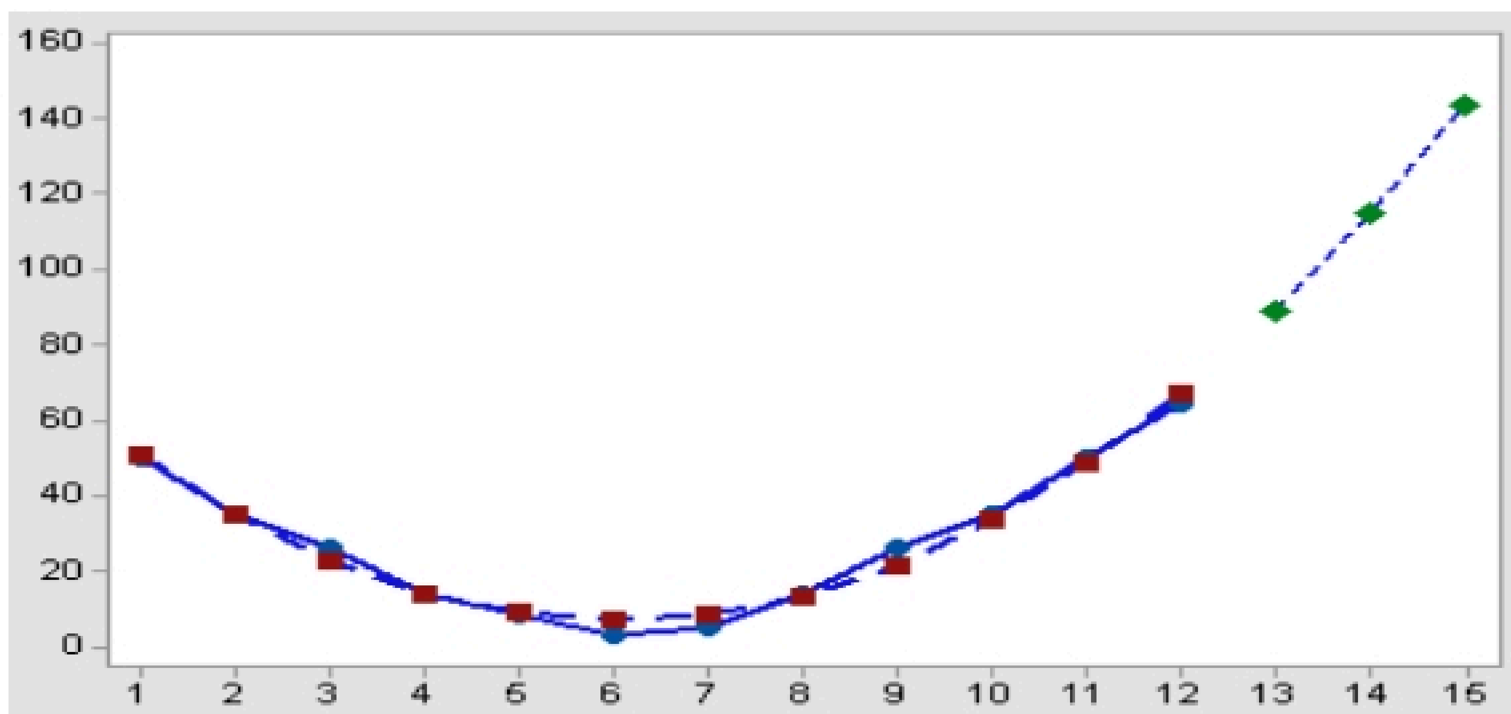
| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 50 | 35 | 26 | 14 | 8 | 3 | 5 | 14 | 26 | 35 | 50 | 64 |
| $t$ | 13 | | 14 | | 15 | | لحظات زمنية مستقبليّة | | | | | |
| $\hat{x}_t$ | 89.22 | | 114.71 | | 143.62 | | القيم المتنبأ بها | | | | | |

وتمثيلها البياني يعرضه الشكل الآتي:




الشكل (١١ ,١٢.د)

وباستخدام برنامج Minitab نجد العرض الآتي للتوفيق والتنبؤات:




الشكل (١١ ,١٢. هـ)

### (١٢,٣,٣) تقدير الاتجاه العام بوساطة دوال النمو Estimate the Trend by Growth Function

من النماذج الشهيرة التي لها دور مهمّ في تقدير مسار بعض المتسلسلات الزمنية ما يُعرف باسم **متسلسلات النمو** Growth Time Series. حيث تُعدّ متسلسلات النمو من النماذج المهمّة التي لها تطبيقات في ميادين علميّة متعدّدة مثل الاقتصاد، والعلوم الطبيعية، والطب و...، وبسبب أهميتها سنقوم باستنتاج نموذج دالَّة النمو من خلال المثالين التطبيقيين الآتيين.

#### (١٢,٣,٣,١) أمثلة

١ – لدى مراقبة سرعة نمو طفل منذ الولادة وحتى عام كامل (قبل دخول مؤثرات خارجية تؤثر في نمو الطفل كتغيير طريقة الطعام والسير و...) وجد أنّ سرعة نمو جسم هذا الطفل مع مرور الزمن $t$ تخضع للمعادلة التفاضلية الآتية:

$$\frac{d\,\mathbf{m}(t)}{dt}=\frac{b}{m}\mathbf{m}(t)(m-\mathbf{m}(t)) \qquad (1)$$

علماً أنَّ $m>0$ يُمثِّل الوزن النهائي للطفل، و $b>0$ يمثِّل ثابت سرعة النمو للطفل، وأمّاَ $\mathbf{m}(t)$ فإنَّه يمثِّل وزن الطفل عند لحظة زمنية معيّنة $t$.

الآن بضرب المعادلة (1) بـ $\frac{1}{\mathbf{m}(t)}$ ينتج لدينا العلاقة الآتية:

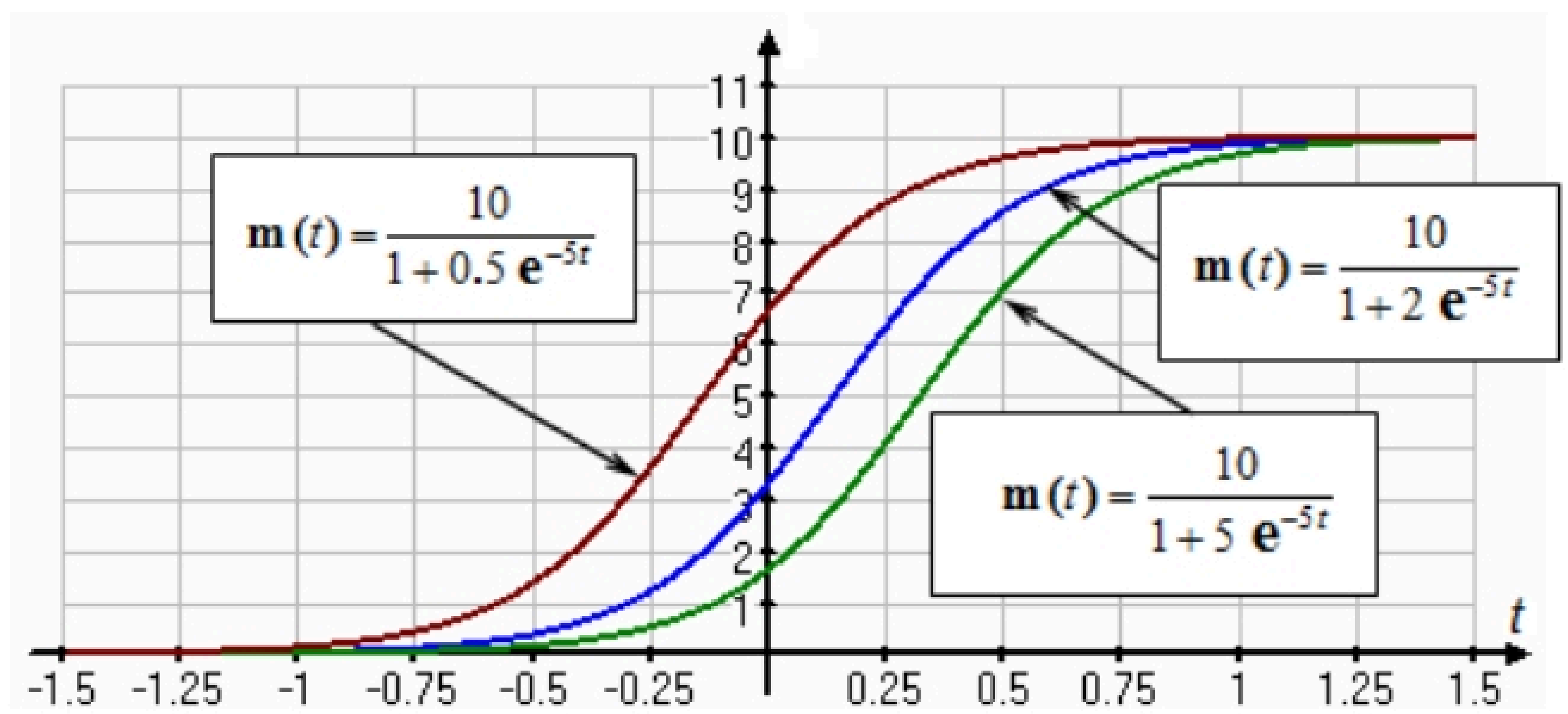
$$\frac{d\,\mathbf{m}(t)}{dt}\frac{1}{\mathbf{m}(t)}=\frac{b}{m}\left(m-\mathbf{m}(t)\right) \qquad (2)$$

حيث نلاحظ أنَّ الطرف الأيسر من العلاقة (2) يمثل السرعة النسبية لنمو الطفل، وهذه العلاقة الأخيرة هي معادلة خطية بـ $\mathbf{m}(t)$ حلّها من الشكل:

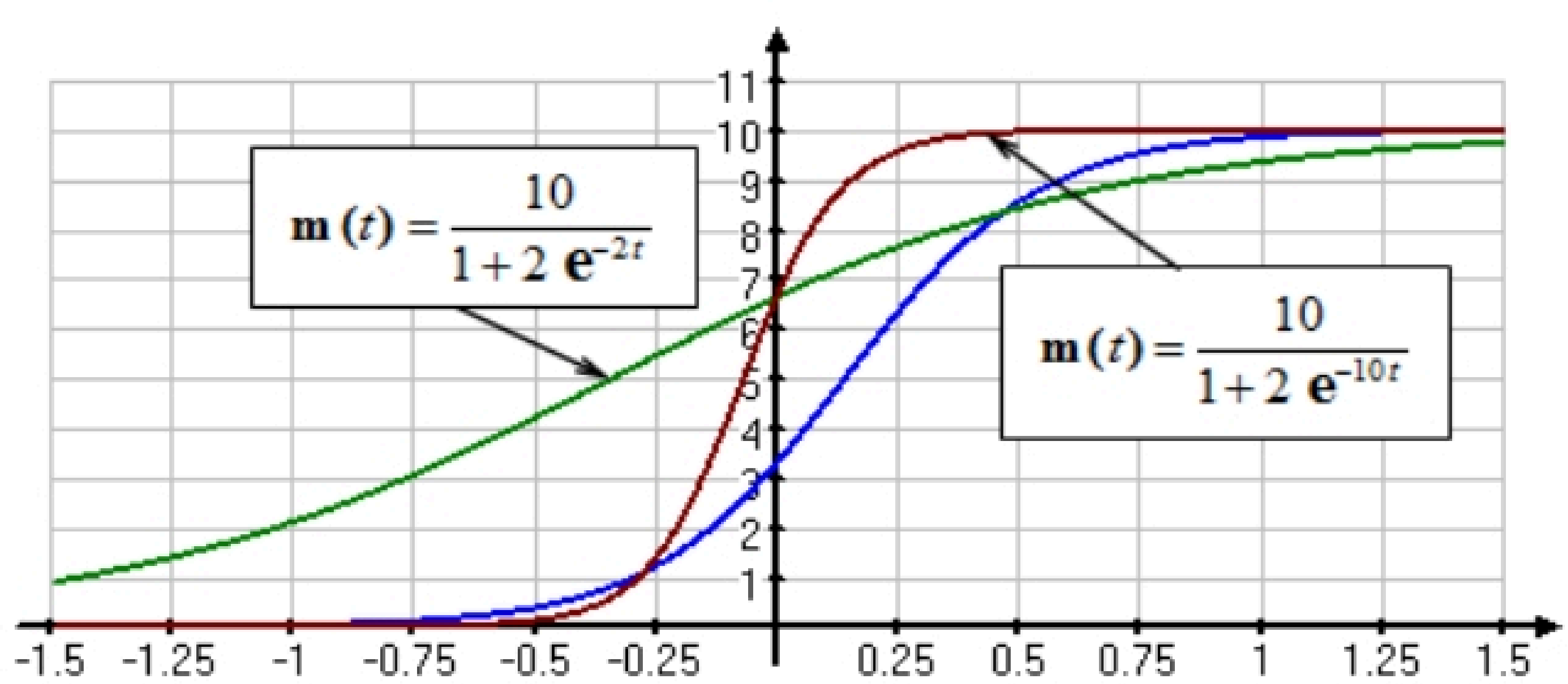
$$\mathbf{m}(t)=\frac{m}{1+a\,e^{-bt}} \qquad (3)$$

علماً أنَّ الثابت $a>0$ ينتج من الشرط الحديّ $\lim\limits_{t\to t_\infty}\mathbf{m}(t)=m$ والذي يوافق القيمة العظمى للسرعة النسبية لنمو جسم الطفل، حيث إنَّ $t_\infty$ ترمز إلى أكبر قيمة ممكنة لـ $t$.

من أجل أخذ تصور أوضح حول سلوك هذه الدالَّة الممثَّلة بالعلاقة (3) سنعرض الرسوم البيانية لهذه الدالَّة من أجل قيمة ثابتة لـ $m$ وقيم متعدّدة لكل من $a$ و$b$، فنجد لها العروض الآتية:




الشكل (١٢, ١٢. أ)




الشكل (١٢, ١٢. ب)

إنّ الدالَّة الممثَّلة بالعلاقة (3) تُعرف باسم **الدالَّة المنطقية** Logistic Function ، وتُسمَّى **دالَّة النمو** Growth Function أيضاً، وهذه المعادلة تحوي ثلاث مَعَالم ولتقديرها نقوم بخطيتها (**بتحويلها إلى معادلة خطية**) بعد أن نفترض أنّ قيمة $m$ معلومة لأنَّها تُحسب من الشرط الحدي $\lim\limits_{t \to t_\infty} \mathbf{m}(t) = m$ ، ولأجل خطية الدالَّة الممثَّلة بالعلاقة (3) نكتب:

$$m = \mathbf{m}(t) + \mathbf{m}(t)\, a\, \mathbf{e}^{-bt}$$

ومنه يكون لدينا:

$$\ln(m - \mathbf{m}(t)) = \ln \mathbf{m}(t) + \ln a - bt$$

وأخيراً ينتج لدينا:

$$\ln \frac{m - \mathbf{m}(t)}{\mathbf{m}(t)} = \ln a - b\,t$$

وبوضع $a^* = \ln a$ و $z_t = \ln \dfrac{m - \mathbf{m}(t)}{\mathbf{m}(t)}$ نحصل على العلاقة الخطية $z_t = a^* - b\, t$ (**خطية بـ $t$**) ، علماً أنَّ قيم $\mathbf{m}(t)$ تعيَّن من القيم المشاهدة $x_t$ وذلك بوضع $\mathbf{m}(t) = x_t$ . بعد ذلك نقوم بتعيين مستقيم المربعات الصغرى للبيانات $(t, z_t)$ عندما نمسح كل القيم الممكنة لـ $t$ ، ومن ثَّم نقوم بالتعويض بشكلٍ عكسي لما افترضاه للخطية فنحصل على معادلة المنحنى المقدِّر للاتجاه العام للبيانات المقدَّمة للمتسلسلة الزمنية.

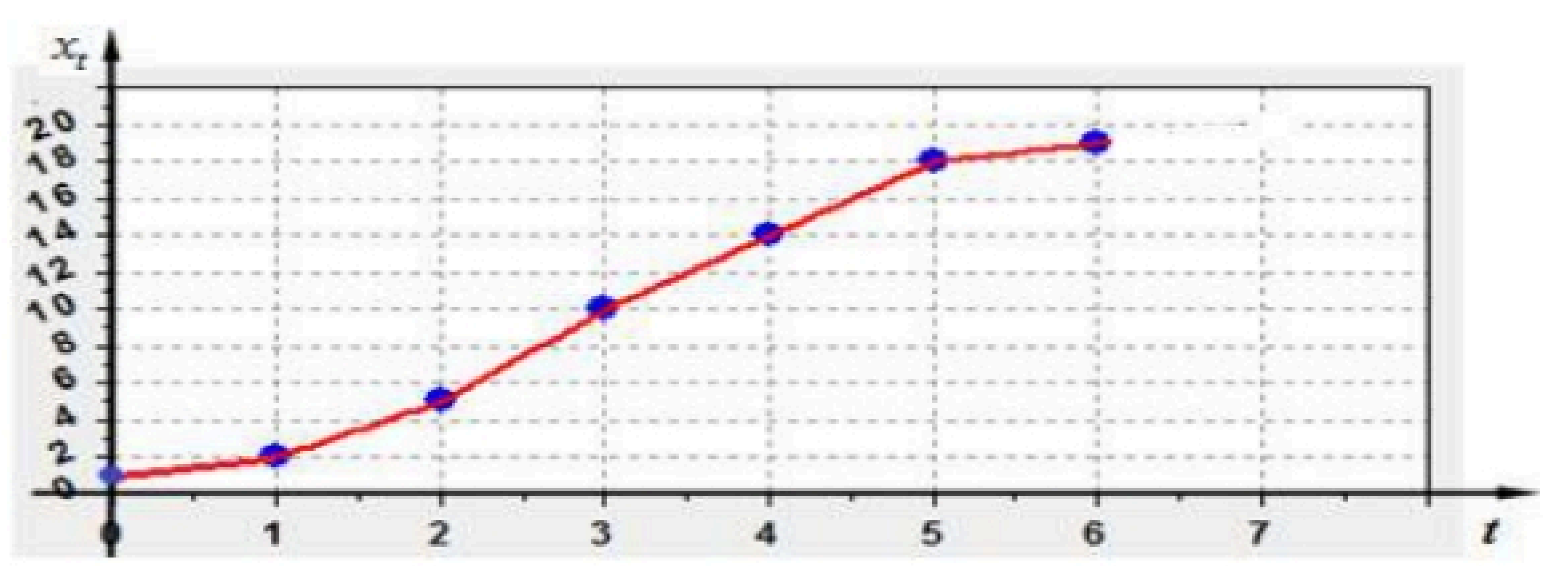
المثال الآتي يُعدّ من الأمثلة التي تنسجم مع تطبيقات دالَّة النمو أيضاً.

٢- لدينا تجربة لقياس سرعة ردود الأفعال عند الأشخاص من خلال مؤثّرات صوتية وضوئية وحركية وبحيث يمكن لأي شخص أن ينال 20 نقطة من 20 نقطة كحد أقصى. أخضع شخص لهذه التجربة وقام بتكرار هذه التجربة على مدى سبعة أيام متتالية مستفيداً من خبرته في كل تكرار لتحصيل نتيجة أفضل في التكرار اللاحق فكانت نتائجه على النحو الآتي:

الجدول (٧, ١٢. أ)

| $t$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $x_t$ | 1 | 2 | 5 | 10 | 14 | 18 | 19 |

أمّا بصمة المتسلسلة الزمنية لنتائج الشخص المتدرب فلها الشكل الآتي:




شكل (١٣, ١٢ أ).

فنلاحظ أن لبصمه هذه المتسلسلة شكل دالَّة نمو، ولذلك سنستخدم النموذج السابق في تقدير الاتجاه العام لهذه المتسلسلة الزمنية.
حيث لدينا في هذا التطبيق $m = \lim\limits_{t \to t_\infty} \mathbf{m}(t) = 20$ وكذلك:

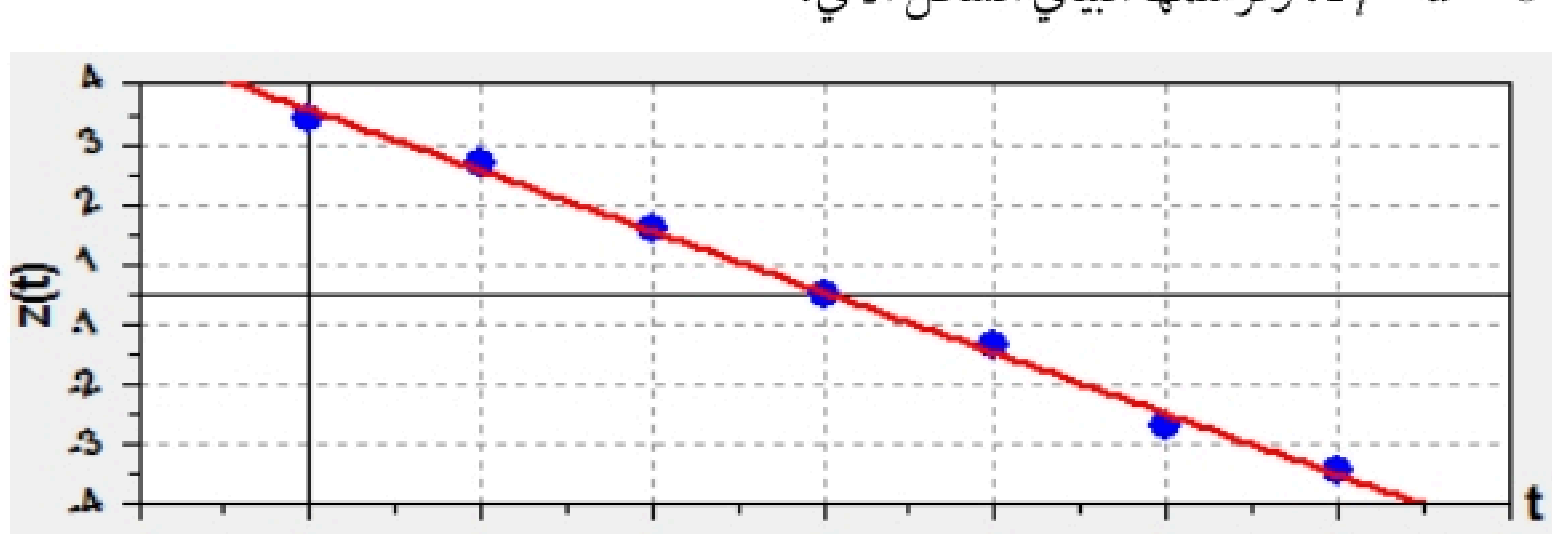
$$z_t = \ln \frac{m - \mathbf{m}(t)}{\mathbf{m}(t)} = \ln \frac{20 - x_t}{x_t}$$

فيكون لـ $z_t$ القيم الموضَّحة في الجدول الآتي:

الجدول (٧, ١٢. ب)

| $t$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $x_t$ | 1 | 2 | 5 | 10 | 14 | 18 | 19 |
| $z_t$ | 2.94 | 2.19 | 1.09 | 0 | -0.84 | -2.19 | -2.94 |

وبتعيين مستقيم المرّبعات الصغرى للبيانات $(t, z_t)$ عندما نمسح كل القيم الممكنة لـ $t$، فإنَّنا سنجد لها المعادلة $z_t = a^* - b\,t = 3.07 - 1.01\,t$، ولرسمها البياني الشكل الآتي:

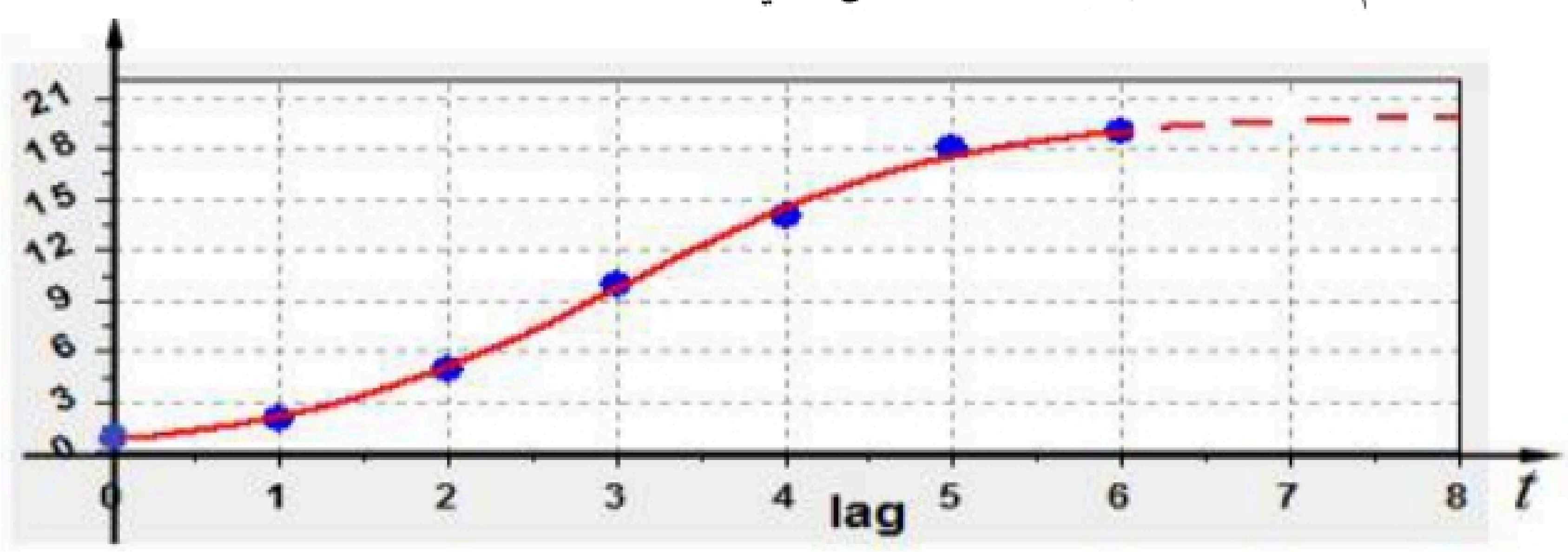



الشكل (١٣, ١٢. ب)

ومن المعادلة السابقة نجد أنَّ $b = -1.01$ و $a = \mathbf{e}^{a^*} = \mathbf{e}^{3.07} = 21.565$، ومن ثَمَّ يصبح للدالَّة المقدِّرة للاتجاه العام للمتسلسلة الزمنية المُعطاة العرض الآتي:

$$\hat{\mathbf{m}}(t) = \frac{m}{1 + a\,\mathbf{e}^{-bt}} = \frac{20}{1 + 21.565\,\mathbf{e}^{-1.01\,t}}$$

ومنه يكون لدالَّة الاتجاه العام للمتسلسلة الزمنية المُعطاة الشكل الآتي:

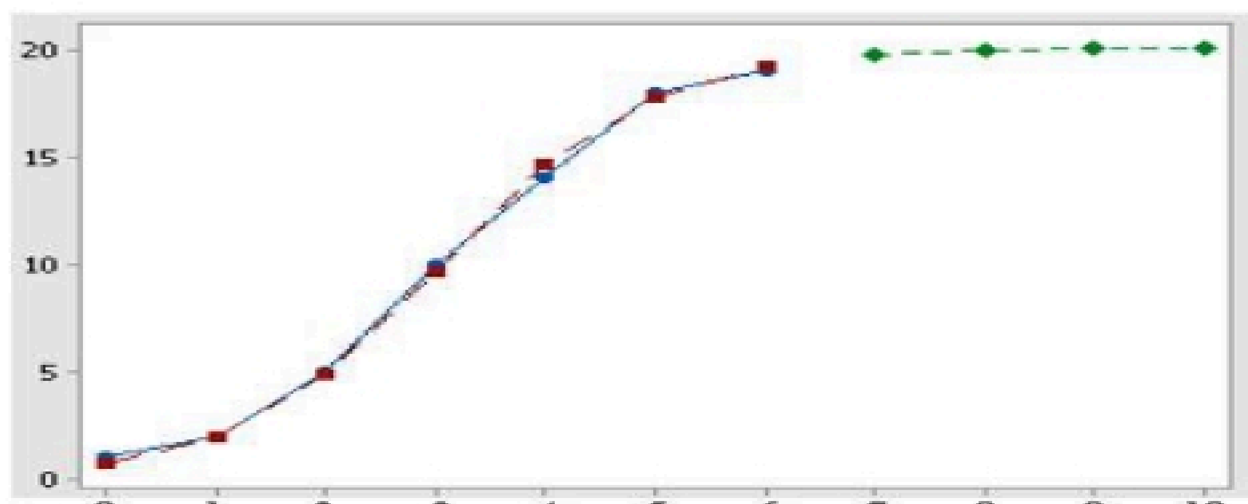



الشكل (١٣, ١٢.ج)

بالطبع كان بإمكاننا تعيين دالَّة الاتجاه العام للمتسلسلة الزمنية المُعطاة باستخدام برنامج Minitab باختيار فقرة تحليل الاتجاه العام Trend، ومن ثَمَّ تفعيل خيار S-Curve فنحصل على العلاقة الآتية كدالَّة مقدِّرة للاتجاه العام للمتسلسلة الزمنية المُعطاة:

$$\hat{\mathbf{m}}(t) = \frac{100}{4.932 + 277.283 \times 0.3752^{\,t}}$$

ولرسمها البياني الشكل الآتي:




الشكل (١٣, ١٢.د)

فنلاحظ أنَّ بيانات المتسلسلة الزمنية قريبة من منحنى الاتجاه العام، ولذلك يُمكننا وبخطأ صغير نسبياً استخدام معادلة منحنى الاتجاه العام $\hat{\mathbf{m}}(t)$ لهذه المتسلسلة الزمنية (ومن أجل أي من النموذجين السابقين) من أجل الحصول على تنبؤات جيدة. حيث نجد التنبؤات للفجوات الزمنية الثلاث التالية 8، 9 و10 مقدَّمة كما في الجدول الآتي:

الجدول (٧, ١٢.هـ)

| الأيام $t$ | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
| 1 | 2 | 5 | 10 | 14 | 18 | 19 | 19.64 | 19.87 | 19.95 | 19.98 |
| البيانات الأصل $x_t$ | | | | | | | قيم متنبأ بها $\hat{x}_t$ | | | |

والرسم البياني لهذه البيانات هو الشكل السابق (١٣ , ١٢ . د) نفسه.

### (١٢،٣،٣،٢) ملاحظات

١- كما لاحظنا، فإنَّه إذا كانت بيانات المتسلسلة الزمنية قريبة جداً من منحنى الاتجاه العام كما في المثالين (١٢،٣،٢،٢) و /٢/ من (١٢،٣،٣،١)، فعندئذ يُمكننا (وبخطأ صغير نسبياً) استخدام معادلة منحنى الاتجاه العام من أجل التنبؤ، حيث يمكننا الحصول على تنبؤات جيدة.

٢- إذا لوحظ أنَّ بيانات المتسلسلة الزمنية تتبعثر بحيث يصعب معها اختيار المنحنى الذي سيمثل الاتجاه العام لهذه المتسلسلة الزمنية؛ فإنَّه يمكن استخدام كثيرات الحدود لتعيين الاتجاه العام، وذلك لأنَّه (وكما ذكرنا سابقاً) من فوائد استخدام كثيرات الحدود رفع درجة كثيرة الحدود للحصول على منحنى أكثر توافقاً مع البيانات المُعطاة. لكن عندما تصبح درجة كثيرة الحدود أكبر من (2) فإنَّ مَعَالِم كثيرة الحدود تصبح صعبة التفسير أو عديمة التفسير، فعلى سبيل المثال إذا كانت كثيرة الحدود من الدرجة الثانية:

$$\mathbf{m}(t) = a + bt + ct^2$$

فعندئذ يمكن تفسير الأمثال $a$ ، $b$ و $c$ على النحو الآتي:

الثابت $a$ يدّل على القيمة الابتدائية للمتسلسلة الزمنية أي إنَّ $a = \hat{\mathbf{m}}(0)$، الثابت $b$ يدّل على سرعة تزايد بيانات المتسلسلة الزمنية، وأمّا الثابت $c$ فإنَّه يدّل على تسارع (العجلة) التزايد لبيانات المتسلسلة الزمنية، فلو كانت كثيرة الحدود من الدرجة الثالثة:

$$\mathbf{m}(t) = a + bt + ct^2 + dt^3$$

فما هو التفسير الموافق لـ $d$ ؟

٣- ذُكر في المرجع [33] أنَّه يمكن للمرء أن يستخدم ناتج طريقة التجزئة للتنبؤ حتى فترة طولها يساوي نصف طول الفترة الزمنية التي استخدمت في الدراسة، ولكن أرى أنَّ هذا الطرح ينطوي على شيء من عدم الدقّة في الحالة العامّة وخاصَّة إذا كانت المتسلسلة الزمنية تحوي مركّبات موسميّة أو دوريّة، فعندئذ التنبؤ لفترة زمنية تزيد عن طول دور كامل ليس له معنى في الواقع العملي.

في بعض الحالات قد يكون اهتمامنا منصبًّا على معرفة إن كان لمتسلسلة زمنية قيد الدراسة اتجاهًا عامًّا مطّردًا أو لا، وبشكل خاص عندما يصعب علينا استنباط الاتجاه العام من خلال النظر إلى بصمة المتسلسلة الزمنية، ومثال على ذلك الشكل (١٢،١٥.أ) في المثال القادم / ٢/ من (١٢،٣،٤،١). في مثل هذه الحالات توجد اختبارات عديدة للبت في مثل هذه المسألة ومنها الاختبار الآتي.

### (١٢،٣،٤) اختبار الفرق لـ نويمان Neumann Deference Test

إنَّ اختبار الفرق لـ نويمان يبيِّن إن كان لمتسلسلة زمنية قيد الدراسة اتجاهًا عامًّا مطّردًا أم لا، وذلك من خلال اختبار الاستقلال لبيانات مأخوذة من مسار متسلسلة زمنية، فلو افترضنا أنَّه لدينا عيِّنة $x_1$ ، $x_2$ ،...و $x_\tau$ مأخوذة من مسار متسلسلة زمنية على فترات زمنية متساوية الطول. عندئذ اختبار الفرق لـ نويمان يقوم على اختبار الفرضية الابتدائية الآتية:

$\mathbf{H}_o$ : البيانات $x_1$ ، $x_2$ ،...و $x_\tau$ هي قيم عشوائية مستقلّة

مقابل الفرضية البديلة:

$\mathbf{H}_A$ : البيانات $x_1$ ، $x_2$ ،...و $x_\tau$ هي قيم عشوائية مرتبطة

وذلك عند مستوى من الأهميّة $\alpha$ .

**من أجل اتخاذ القرار بشأن هذا الاختبار نقوم بما يلي:**

١- نحسب الفروق المتتالية الأولى:

$$\Delta x_t = x_{t+1} - x_t \qquad ; t = 1,2,3,\dots, \tau - 1$$

٢- نحسب المتوسطات لمربّعات الفروقات السابقة:

$$\overline{\Delta x_t^{\,2}} = \frac{1}{\tau-1}\sum_{t=1}^{\tau-1}\Delta x_t^{\,2} \quad ; t = 1,2,3,\dots, \tau-1$$

٣- نقوم بحساب قيمة الإحصاءة:

$$\mathbf{NE} := \frac{\overline{\Delta x_t^{\,2}}}{s^2} = \overline{\Delta x_t^{\,2}} \cdot \left(\frac{1}{\tau-1}\sum_{t=1}^{\tau}(x_t - \bar{x})^2\right)^{-1} \qquad [12,4]$$

علماً أنَّ $s$ هو الانحراف المعياري للبيانات $x_1$، $x_2$،...، و $x_\tau$.

٤- إذا كان:

أ- حجم العيِّنة $\tau > 60$ فإنَّنا نضع:

$$q_{1-\alpha} = 2 - 2\, z_{1-\alpha}\sqrt{\frac{\tau-2}{(\tau-1)(\tau+1)}} \qquad [12,5]$$

علماً أنَّ $z_{1-\alpha}$ هي القيمة التي تمَّ استخدامها في الفصول الثلاث السابقة وتستخرج من جدول التوزيع الطبيعي المعياري.

ب- حجم العيِّنة $\tau \leq 60$ فإنَّنا نأخذ قيمة $q_{1-\alpha}$ من الجدول الآتي والخاص بهذا الاختبار.

الجدول (١٢,٨)

| $\tau$ | | 10 | 15 | 20 | 25 | 30 | 35 | 40 | 45 | 50 | 55 | 60 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $\alpha = 0.05$ | $q_{1-\alpha}$ | 1.06 | 1.21 | 1.30 | 1.37 | 1.41 | 1.46 | 1.49 | 1.52 | 1.54 | 1.56 | 1.58 |
| $\alpha = 0.01$ | $q_{1-\alpha}$ | 0.75 | 0.92 | 1.04 | 1.13 | 1.20 | 1.25 | 1.29 | 1.33 | 1.36 | 1.39 | 1.41 |

٥- **اتخاذ القرار**: اذا كان $\mathbf{NE} < q_{1-\alpha}$ فعندئذ تُرفض الفرضية الابتدائية $\mathbf{H}_o$ عند مستوى الأهمّيّة $\alpha$، وهذا يعني أنَّ البيانات المقدَّمة مرتبطة عند مستوى الأهمّيّة $\alpha$، ومن ثمَّ المتسلسلة الزمنية تملك اتجاهًا عامًّا مطّردًا عند هذا المستوى من الأهمّيّة.

بالطبع توجد اختبارات أخرى لتحديد إن كان لمتسلسلة زمنية اتجاهًا عامًّا مطّردًا أم لا، ومنها Moore–Wallis Test، وكذلك اختبار Staurt–Cox Test (انظر [33]).

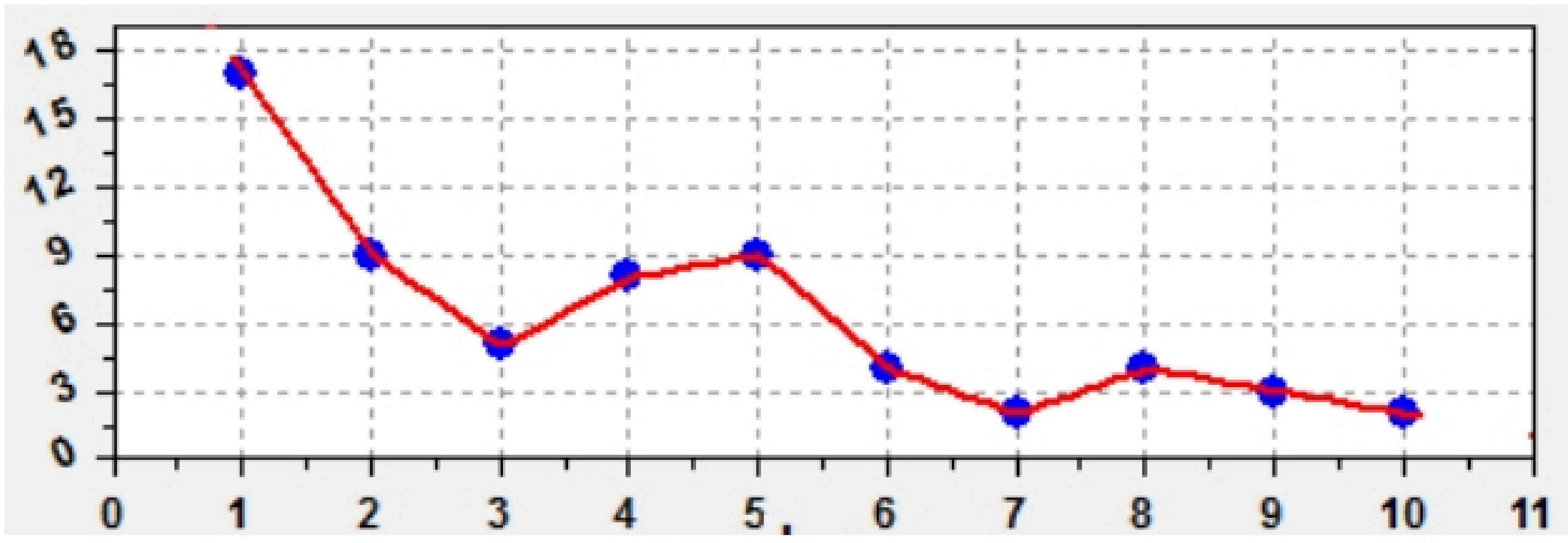
### (١٢,٣,٤,١) أمثلة

١- لتكن لدينا المشاهدات المقدَّمة في الجدول الآتي والمأخوذة من مسار متسلسلة زمنية تصف عدد الأخطاء المرتكبة في مباريات فريق لكرة القدم منذ بداية الدوري العام السّنوي.

الجدول (١٢,٩ .أ)

| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 17 | 9 | 5 | 8 | 9 | 4 | 2 | 4 | 3 | 2 |

فنجد لبصمة هذه المتسلسلة الزمنية العرض الآتي:




الشكل (١٤ , ١٢)

الآن لنختبر إن كانت هذه البيانات مستقلّةً بعضها عن بعض أم لا (بمعنى إن كان لهذه البيانات اتجاه عام مطّرد أم لا) وذلك عند مستوى من الأهمّية $\alpha = 0.05$ .

الحـل: من أجل ذلك علينا اختبار الفرضية الابتدائية:

$\mathbf{H}_o$ : البيانات $x_1$ ، $x_2$ ،...و $x_T$ هي قيم عشوائية مستقلّة

مقابل الفرضية البديلة:

$\mathbf{H}_A$ : البيانات $x_1$ ، $x_2$ ،...و $x_T$ هي قيم عشوائية مرتبطة

وبحساب الفروق المتتالية الأولى (انظر الجدول (٩ , ١٢ . ب) للفروق المتتالية الأولى)، ومن ثمَّ حساب المتوسطات لمربّعات الفروقات السّابقة ، فإنَّنا سنجد $\overline{\Delta x_t^2} = 13.89$ و $s^2 = 21.34$ ، ومن ثمَّ تكون قيمة الإحصاءة NE تساوي:

$$\mathbf{NE} := \frac{\overline{\Delta x_t^2}}{s^2} = \frac{13.89}{21.34} = 0.65$$

الجدول (٩ , ١٢ . ب)

| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 17 | 9 | 5 | 8 | 9 | 4 | 2 | 4 | 3 | 2 |
| $\Delta x_t$ | -8 | -4 | 3 | 1 | -5 | -2 | 2 | -1 | -1 | |

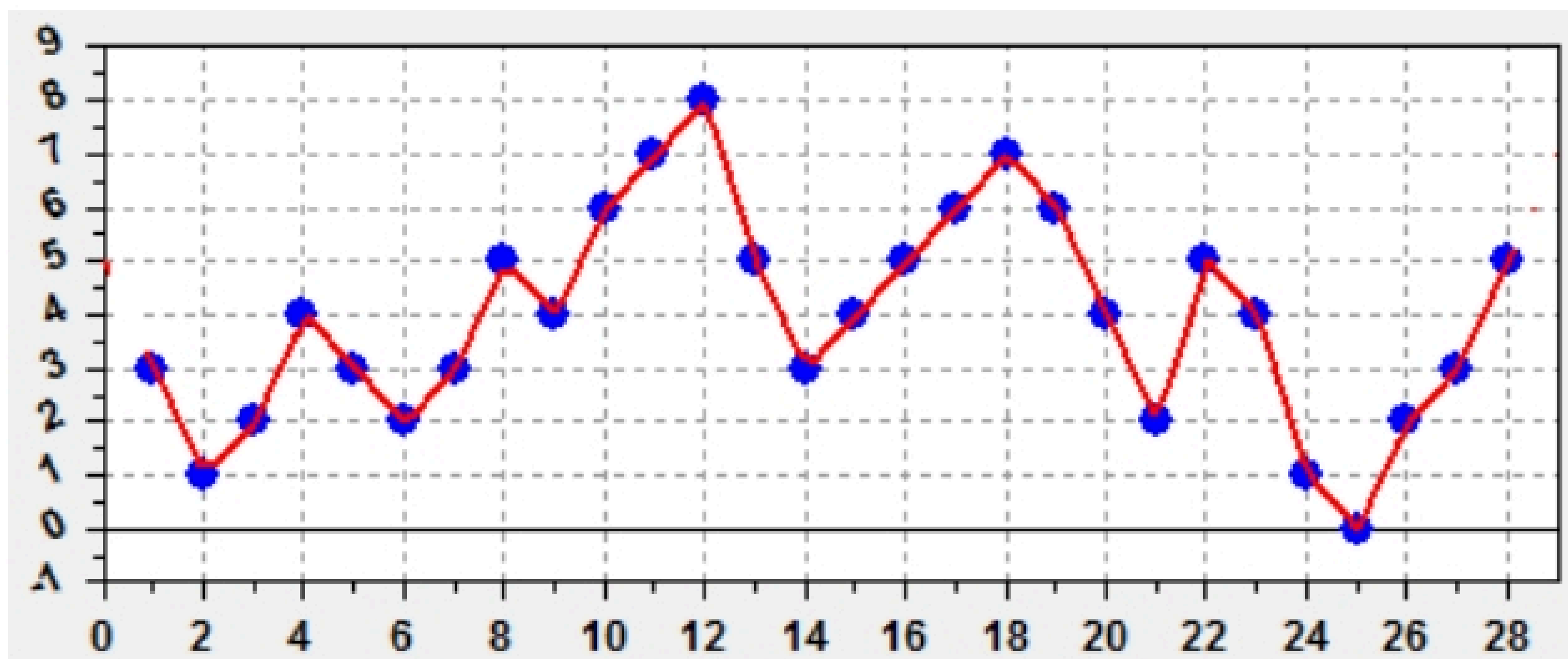
وبما أنَّ حجم العيِّنة $n = 10$ ، ومستوى الأهمّية $\alpha = 0.05$ فإنَّه سيكون لدينا $q_{1-\alpha} = q_{0.95} = 1.06$ (من الجدول (٨ , ١٢ ))، وهكذا نجد أنَّ $\mathbf{NE} < q_{1-\alpha}$ ، ومن ثمَّ نرفض الفرضية الابتدائية $\mathbf{H}_o$ عند مستوى من الأهمّية $\alpha = 0.05$ ، ومن ثمّ فإنَّ عدد الأخطاء المرتكبة في مباريات هذا الفريق ليست مستقلَّةً بعضها عن بعض، وبالتالي المتسلسلة الزمنيةَ لظاهرة عدد الأخطاء المرتكبة في مباريات هذا الفريق تملك اتجاهاً عاماً مطّرداً (وهو اتجاه عام هابط كما أظهره الشكل (١٤ , ١٢ )) عند مستوى الأهمّية $\alpha = 0.05$ .

٢ – لتكن لدينا المشاهدات المقدَّمة في الجدول الآتي والمأخوذة من مسار متسلسلة زمنية تصف عدد نوبات الصرع في كل شهر لدى مريض وعلى مدى 28 شهراً، ولنبحث هل كان عدد النوبات الشهريّة للمريض مستقلَّة بعضها عن بعض أم لا؟

الجدول (١٠ , ١٢ . أ)

| الشهر $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 3 | 1 | 2 | 4 | 3 | 2 | 3 | 5 | 6 | 4 | 7 | 8 | 5 | 3 |
| الشهر $t$ | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| $x_t$ | 4 | 5 | 6 | 7 | 6 | 4 | 2 | 5 | 4 | 1 | 0 | 2 | 3 | 5 |

فنجد لبصمة هذه المتسلسلة الزمنية العرض الآتي:




شكل (١٥ , ١٢ أ).

الآن لنختبر إن كان للمتسلسلة الزمنية المُعطاة اتجاهًا عامًّا مطّردًا أم لا (أي إن كانت هذه البيانات مستقلة بعضها عن بعض أم لا)، وذلك عند مستوى من الأهمّيّة قدره $\alpha = 0.05$ .

الحـل: من أجل ذلك لنختبر الفرضية الابتدائية:

$\mathbf{H}_o$ : البيانات $x_1$ ، $x_2$ ،...و $x_\tau$ هي قيم عشوائية مستقلّة

مقابل الفرضية البديلة:

$\mathbf{H}_A$ : البيانات $x_1$ ، $x_2$ ،...و $x_\tau$ هي قيم عشوائية مرتبطة

وبحساب الفروق المتتالية الأولى، ومن ثمَّ حساب المتوسطات لمربّعات الفروقات السابقة نجد ما يلي (انظر الجدول الآتي للفروق المتتالية الأولى):

الجدول (١٠ , ١٢ . ب)

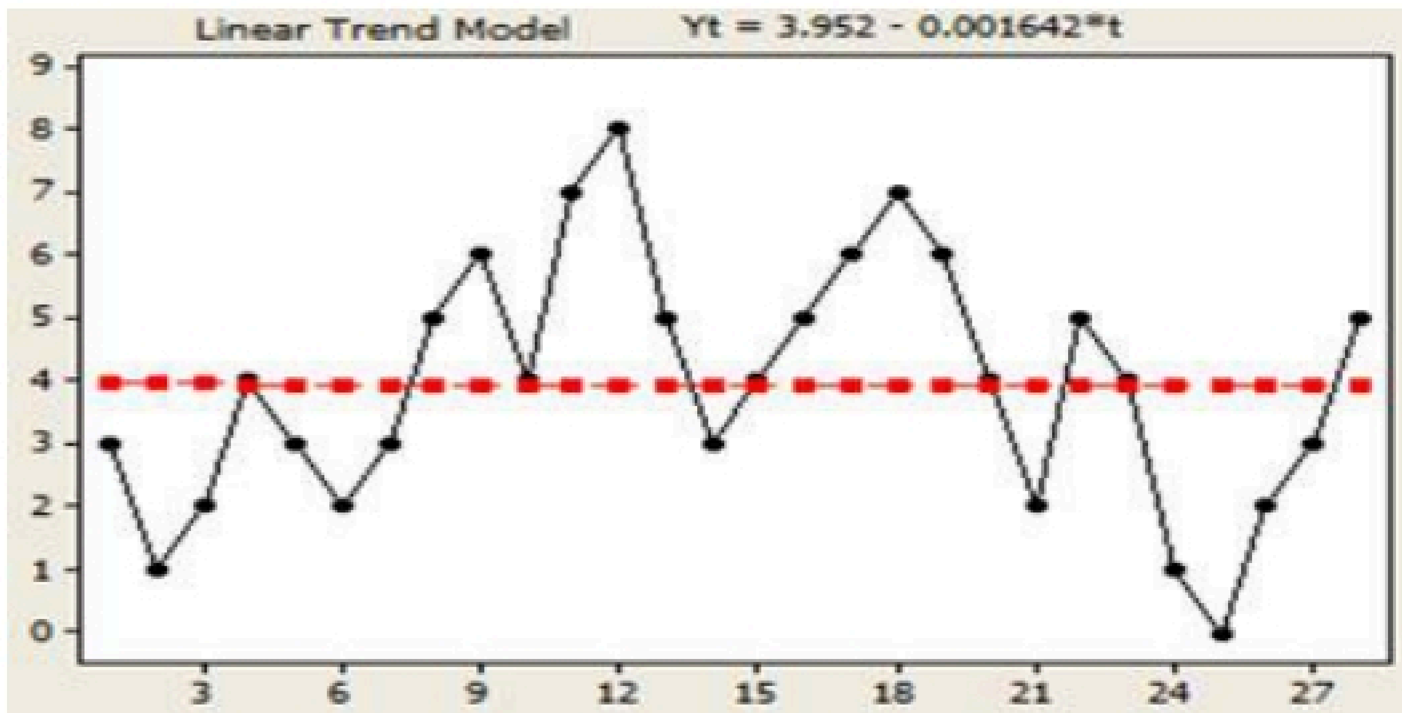
| الشهر $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 3 | 1 | 2 | 4 | 3 | 2 | 3 | 5 | 6 | 4 | 7 | 8 | 5 | 3 |
| $\Delta x_t$ | | -2 | 1 | 2 | -1 | -1 | 1 | 2 | 1 | -2 | 3 | 1 | -3 | -2 |
| الشهر $t$ | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| $x_t$ | 4 | 5 | 6 | 7 | 6 | 4 | 2 | 5 | 4 | 1 | 0 | 2 | 3 | 5 |
| $\Delta x_t$ | | 1 | 1 | 1 | -1 | -2 | -2 | 3 | -1 | -3 | -1 | 2 | 1 | 2 |

فيكون لدينا $\overline{\Delta x_t^{\,2}} = 3.19$ و $s^2 = 3.921$ ، ومنه نجد أنَّ قيمة الإحصاءة $\mathbf{NE}$ تساوي:

$$\mathbf{NE} := \frac{\overline{\Delta x_t^{\,2}}}{s^2} = \frac{3.19}{3.92} = 0.814$$

وبما أنَّ حجم العيِّنة $n = 28$ ، ومستوى الأهمّيّة $\alpha = 0.05$ فإنَّه سيكون لدينا $q_{0.95} = 1.41$ (من الجدول (٨ , ١٢ ))، وهكذا ينتج لدينا أنَّ $\mathbf{NE} > q_{1-\alpha}$، وبالتالي نرفض الفرضية الابتدائية $\mathbf{H}_o$ عند مستوى من الأهمّيّة $\alpha = 0.05$ ، ومن ثَمّ فإنَّ عدد نوبات الصرع الشـــهريّة للمريض ليســت مســـتقلة بعضها عن بعض (مرتبطة بعضها ببعض)، وهكذا نجد أنَّ المتسلسلة الزمنيةَ لظاهرة الصرع لدى المريض تملك اتجاهاً عاماً مطّرداً عند مســتوى الأهمّيّة $\alpha = 0.05$ . لاحـظ صعوبة تحديد اطراد الاتجاه العام (الممثَّل بمستقيم الانحدار)

لهذه المتسلسلة الزمنية بالعين المجردة؛ وذلك لأنَّ ميل مستقيم الانحدار لهذه البيانات صغير جداً ويساوي 0.001642– (انظر الشكل الآتي).




الشكل (١٥ ,١٢. ب)

## (١٢,٤) دراسة المركبة الدورية لمتسلسلة زمنية

**Study the Period Component of Time Series**

الآن وبعد التعرّف على مفهوم ومدلول الاتجاه العام لمتسلسلة زمنية وكيفية تعيينه سننتقل إلى دراسة المُركَّبة الدوريّة. في الواقع توجد طرائق عديدة لتقدير المركِّبة الدوريّة لمتسلسلة زمنية، وكل طريقة تتطلب شروطاً محدَّدة لتطبيقها، ومن هذه الطرائق (على سبيل المثال لا الحصر):

أ- طريقة جيب التمام (أو التجيب) لـ هلبرغ.

ب- طريقة تحليل الارتباط والترددات الخفية.

ج- طريقة العارض الدوري.

د- طريقة تقدير الأدوار لدى أدوار بأطوال مختلفة.

في كتابنا هذا سنقتصر على تقديم الطريقة الأولى فقط، وهي تمتاز ببساطتها ووضوحها، ولكن لها شروط مُحَدَّدة لاستخدامها.

### (١ , ٤ , ١٢) طريقة جيب التمام (أو التجيب) لـ هلبرغ Halberg's Cosine Method

لقد قدَّم الرياضياتي الفرنسي هلبرغ (1919-2013) Franz Halberg طريقة لتقدير المركِّبة الموسميّة لمتسلسلة زمنية تحتوي في مسارها على اهتزازات توافقية (جيبية أو قريبة من الجيبية)، وهذه الطريقة تنطلق من نموذج الجمع الآتي:

$$X_t = \mathbf{m}(t) + \mathbf{s}(t) + N_t \qquad [12,6]$$

والمركِّبة الموسميّة $\mathbf{s}(t)$ هنا تتضمَّن كل من المركِّبة الدوريّة والموسميّة $\mathbf{p}(t)+\mathbf{s}(t)$ في نموذج التجزئة ذي المركِّبات الأربع، ويقوم هذا الاختبار على شروط عدّة يجب توفرها، وهي:

أ- المتسلسلة الزمنية تحتوي على اهتزازات توافقية.

ب- أعداد القياسات في كل دور متساوية على وجه التقريب، ويجب ألا يقل عن أربعة قياسات في كل دور.

عندئذ نضع:

$$\mathbf{s}(t) = m + \alpha \, \mathbf{cos}(\omega t + \phi) \qquad [12,7]$$

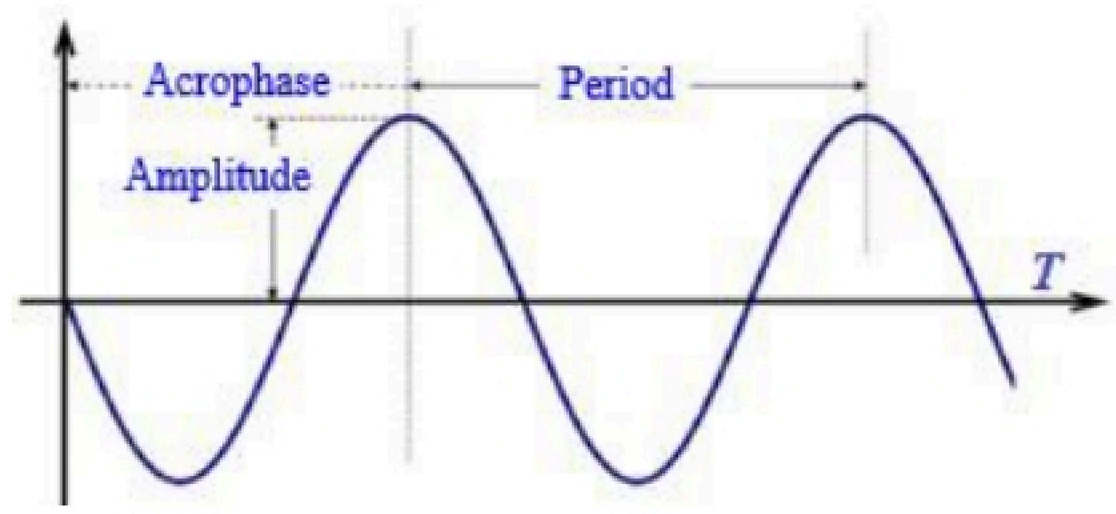
علماً أنَّ:

$m$ هي القيمة الوسطى للمتسلسلة الزمنية.

$\alpha$ هو سعة مطال Amplitude الاهتزاز.

$\omega$ السرعة الزاوية للاهتزاز Angular Frequency التي تساوي $\omega = \frac{2\pi}{p}$، علماً أنَّ $p$ هو طول الدور Length of Period.

$\phi$ يمثِّل فرق الصفحة Acrophase بين المتسلسلة الزمنية ودالَّة جيب التمام (انظر الشكل الآتي).




الشكل (١٦ ,١٢)

وهكذا نلاحظ أنَّه تحت الفرض أنَّ طول الدور معلوماً (ومن ثَمَّ $\omega$ معلوماً دوماً) فإنَّه لدينا ثلاثة مَعَالم يجب تقديرها وهي $m$، $\alpha$ و $\phi$، ولتقدير هذه المَعَالم نقوم بما يلي:

١- نقوم بتقدير مُركِّبة الاتجاه العام $\mathbf{m}(t)$، ومن ثَمَّ نزيلها من الطرف الأيمن لنموذج المتسلسلة الزمنية من خلال وضع:

$$\tilde{X}_t := X_t - \hat{\mathbf{m}}(t) = \mathbf{s}(t) + N_t \qquad [12,8]$$

وفي هذه الحالة يُقال عن النموذج [12,8] للمتسلسلة الزمنية إنَّه معدَّل اتجاهياً، أو إنَّ المتسلسلة الزمنية مقدَّمة بالعرض المُخْتزل.

٢- نفترض أنَّ للدالَّة $\mathbf{s}(t)$ العرض $\mathbf{s}(t) = \alpha \cos(\omega t + \phi)$. عندئذ يمكننا أن نكتب:

$$\tilde{X}_t = m + \alpha \cos(\omega t + \phi) + N_t = m + \alpha \cos(\phi) \cos(\omega t) - \alpha \sin(\phi) \sin(\omega t) + N_t$$

٣- نضع:

$$\beta = \alpha \cos\phi \quad \& \quad \gamma = -\alpha \sin\phi \quad \& \quad u(t) = \cos(\omega t) \quad \& \quad z(t) = \sin(\omega t)$$

وبالتعويض يصبح لدينا:

$$\tilde{X}_t = m + \beta \cdot u(t) + \gamma \cdot z(t) + N_t$$

وهي علاقة خطية بكلٍّ من $z(t)$ و $u(t)$ يُقدَّر فيها $\beta$ و $\gamma$ بطريقة المرّبعات الصغرى، فإذا كانت العيِّنة المأخوذة من مسار المتسلسلة الزمنية هي $x_1$، $x_2$، ...،و $x_\tau$، فعندئذ تحت هذه المعطيات نضع:

$$\tilde{x}_t = x_t - \hat{\mathbf{m}}(t) \quad ; t = 1, 2, \ldots, \tau$$

والتي تُعرف باسم البواقي Residual النّاتجة عن الاتجاه العام المُقدَّر، وفي هذه الحالة يجب أن يكون المجموع الآتي أصغرياً:

$$\sum_{t=1}^{\tau} \left[ \tilde{x}_t - \left(m + \beta \cdot u(t) + \gamma \cdot z(t)\right)\right]^2$$

وهذا يعني أنَّ:

$$\left( \sum_{t=1}^{\tau} \left[ \tilde{x}_t - \left(m + \beta \cdot u(t) + \gamma \cdot z(t)\right)\right]^2 \right)' = 0$$

ومن هذه العلاقة الأخيرة ينتج لدينا نظام المعادلات الخطية الآتية:

$$\sum_{t=1}^{\tau}\tilde{x}_t = m\cdot\tau+\beta\sum_{t=1}^{\tau}u(t)+\gamma\sum_{t=1}^{\tau}z(t)$$

$$\sum_{t=1}^{\tau}\tilde{x}_t\cdot u(t) = m\cdot\sum_{t=1}^{\tau}u(t)+\beta\sum_{t=1}^{\tau}u^2(t)+\gamma\sum_{t=1}^{\tau}u(t)\cdot z(t)$$

$$\sum_{t=1}^{\tau}\tilde{x}_t\cdot z(t) = m\cdot\sum_{t=1}^{\tau}z(t)+\beta\sum_{t=1}^{\tau}u(t)\cdot z(t)+\gamma\sum_{t=1}^{\tau}z^2(t)$$

والتي يمكن تقديمها بالشكل المصفوفي الآتي أيضاً:

$$\begin{pmatrix} \tau & \sum_{t=1}^{\tau}u(t) & \sum_{t=1}^{\tau}z(t) \\ \sum_{t=1}^{\tau}u(t) & \sum_{t=1}^{\tau}u^2(t) & \sum_{t=1}^{\tau}u(t)z(t) \\ \sum_{t=1}^{\tau}z(t) & \sum_{t=1}^{\tau}u(t)z(t) & \sum_{t=1}^{\tau}z^2(t) \end{pmatrix}\cdot\begin{pmatrix} m \\ \beta \\ \gamma \end{pmatrix}=\begin{pmatrix} \sum_{t=1}^{\tau}\tilde{x}_t \\ \sum_{t=1}^{\tau}\tilde{x}_{t_i}\cdot u(t) \\ \sum_{t=1}^{\tau}\tilde{x}_{t_i}\cdot z(t) \end{pmatrix}$$

وبالحل المشترك لنظام المعادلات الخطية السابقة نحصل على تقديرات المَعالم $m$ ، $\beta$ و $\gamma$ من خلال العلاقات الآتية على الترتيب:

$$\hat{m}=\frac{1}{\tau}\sum_{t=1}^{\tau}\tilde{x}_t \qquad [12,9\text{-a}]$$

$$\hat{\beta}=\frac{2}{\tau}\sum_{t=1}^{\tau}\tilde{x}_t\cdot u(t)=\frac{2}{n}\sum_{t=1}^{\tau}\tilde{x}_t\cdot\mathbf{cos}(\omega t) \qquad [12,9\text{-b}]$$

$$\hat{\gamma}=\frac{2}{\tau}\sum_{t=1}^{\tau}\tilde{x}_t\cdot z(t)=\frac{2}{\tau}\sum_{t=1}^{\tau}\tilde{x}_t\cdot\mathbf{sin}(\omega t) \qquad [12,9\text{-c}]$$

وأمَّا قيمة $\alpha$ فإنَّها تقدَّر بوساطة العلاقة الآتية:

$$\hat{\alpha}=\sqrt{\hat{\beta}^2+\hat{\gamma}^2} \qquad [12,9\text{-d}]$$

في حين أنَّ فرق الصفحة $\phi$ يُقدَّر باستخدام العلاقة الآتية:

$$\hat{\phi}=k\cdot\pi\pm\hat{\theta} \qquad ;k\in Q \qquad [12,9\text{-e}]$$

علماً أنَّ $\hat{\theta}$ مقدّرةً بـ الراديان وتحسب من العلاقة الآتية:

$$\hat{\theta}=\mathbf{arctan}\left|\frac{-\hat{\gamma}}{\hat{\beta}}\right|=\mathbf{arctan}\left|\frac{\hat{\gamma}}{\hat{\beta}}\right| \qquad [12,9\text{-f}]$$

وتُعيَّن إشارة $\hat{\theta}$ وقيمة $k$ من الجدول الآتي الذي يعتمد على إشارتي $\hat{\beta}$ و $\hat{\gamma}$ معاً:

الجدول (١٢,١١)

| | $\hat{\beta} > 0$ | $\hat{\beta} = 0$ | $\hat{\beta} < 0$ |
|---|---|---|---|
| $\hat{\gamma} > 0$ | $\hat{\phi} = -\hat{\theta}$ | $\hat{\phi} = -\hat{\theta}$ | $\hat{\phi} = \hat{\theta} + \pi$ |
| $\hat{\gamma} = 0$ | $\hat{\phi} = \hat{\theta} - 2\pi$ | لا توجد مركّبة دوريّة | $\hat{\phi} = \hat{\theta} - \pi$ |
| $\hat{\gamma} < 0$ | $\hat{\phi} = \hat{\theta} - 2\pi$ | $\hat{\phi} = -\hat{\theta} - \pi$ | $\hat{\phi} = -\hat{\theta} - \pi$ |

بعد ذلك نعوِّض في صيغة المركِّبة الموسميّة وبعد ذلك في صيغة العلاقة التي تعطي نموذج المتسلسلة الزمنية أيضاً فنحصل على النموذج الرياضياتي المقدَّر لتلك المتسلسلة الزمنية.

**(١, ١, ٤, ١٢) مثال**

بالعودة إلى المثال (١, ١, ٣, ١٢) حيث لدينا متسلسلة زمنية مولَّدة بوساطة طوري عشوائي $\{X_t \,; t \geq 0\}$ مُعطى بالعلاقة الآتية:

$$X_t = 0.1\, t + \sin(2\pi / 6) t + N_t \qquad ; t \geq 0$$

فلو أمعنا النظر في بصمة العيِّنة المأخوذة من مسار هذه المتسلسلة الزمنية (انظر الشكل (١٠, ١٢)) نلاحظ وجود ثلاثة أدوار، وطول الدور $p$ يساوي 6 فجوات زمنية بسبب وجود ست قيم مقيسة تقريباً في كل دور، وهذا يعني أنَّ الشروط المطلوبة لتطبيق طريقة جيب التمام لـ هلبرغ مُحقَّقةً، ومن ثمَّ بإمكاننا استخدام طريقة جيب التمام لـ هلبرغ لتقدير المركِّبة الدورية.

لقد قمنا سابقاً بتقدير مركِّبة الاتجاه العام لهذه المتسلسلة الزمنية بوساطة الانحدار الخطي حيث كانت لدينا الدالَّة الخطية المقدَّرة لمُركِّبة الاتجاه العام هي:

$$\mathbf{m}(t) = 0.025 + 0.0909\, t$$

ومن ثمَّ لإزالة الاتجاه العام من هذه المتسلسلة الزمنية نقوم بطرح القيم المقدَّرة $\hat{\mathbf{m}}(t)$ من القيم الأصل $x_t$ وذلك من أجل كلّ القيم الممكنة لـ $t$، فنحصل على القيم $\tilde{x}_t$. أي إنَّه لدينا:

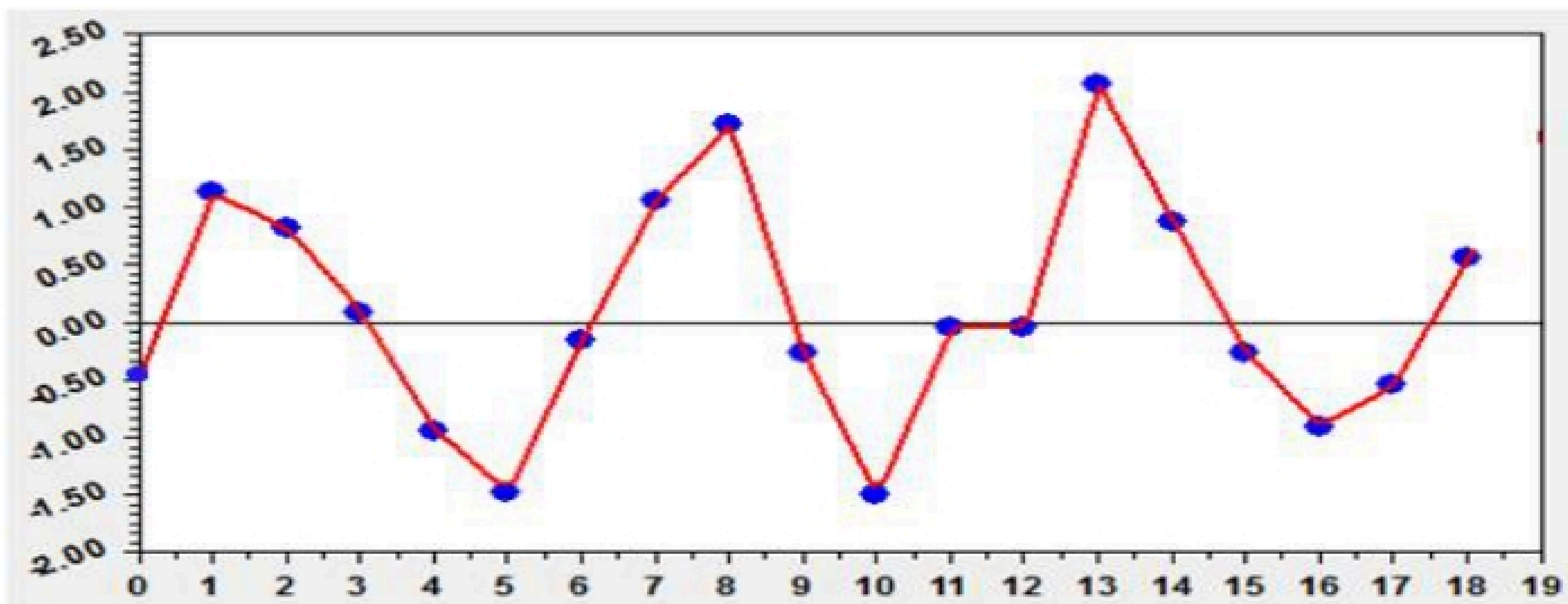
$$\tilde{x}_t = x_t - \hat{\mathbf{m}}(t) \qquad ; t = 0, 1, \ldots, 18$$

ومن أجل المتسلسلة الزمنية التي قيد الدراسة سيكون لـ $\tilde{x}_t$ القيم الآتية:

الجدول (١٢, ١٢. أ)

| $t$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| $\tilde{x}_t$ | -0.475 | 1.1241 | 0.8132 | 0.0923 | -0.9386 | -1.4695 | -0.1504 |
| $t$ | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| $\tilde{x}_t$ | 1.0587 | 1.7178 | -0.2631 | -1.504 | -0.0449 | -0.0358 | 2.0733 |
| $t$ | 14 | 15 | 16 | 17 | 18 | | |
| $\tilde{x}_t$ | 0.8824 | -0.2685 | -0.8994 | -0.5403 | 0.5588 | | |

ولبصمتها الشكل (١٧, ١٢. أ). لاحظ كيف أصبح اتجاهها العام أفقياً وهذا ما يُعبَّر به عادة بإزالة الاتجاه العام لمتسلسلة زمنية.




الشكل (١٧ , ١٢ . أ)

وكذلك باستخدام العلاقات [12,9-a] وحتى [12,9-c] نجد الآتي:

$$\hat{\mathbf{m}} = \frac{1}{19}\sum_{t=0}^{18} \tilde{x}_t = \frac{1}{19}(1.7311) = 0.0911$$

والآن لتقدير المَعْلمتين $\beta$ و $\gamma$ سنبني الجدول الآتي:

الجدول (١٢ , ١٢ . ب)

| $t$ | $x_t$ | $\tilde{x}_t$ | $\tilde{x}_t \cdot u(t)$ | $\tilde{x}_t \cdot z(t)$ |
|---|---|---|---|---|
| 0 | -0.45 | -0.475 | -0.475 | 0 |
| 1 | 1.24 | 1.1241 | 0.5622 | 0.9734 |
| 2 | 1.02 | 0.8132 | -0.4063 | 0.7044 |
| 3 | 0.39 | 0.0923 | -0.0923 | 5E-05 |
| 4 | -0.55 | -0.9386 | 0.4699 | 0.8125 |
| 5 | -0.99 | -1.4695 | -0.7335 | 1.2733 |
| 6 | 0.42 | -0.1504 | -0.1504 | 0.0002 |
| 7 | 1.72 | 1.0587 | 0.5306 | 0.9161 |
| 8 | 2.47 | 1.7178 | -0.8565 | 1.489 |
| 9 | 0.58 | -0.2631 | 0.2631 | -5E-04 |

| $t$ | $x_t$ | $\tilde{x}_t$ | $\tilde{x}_t \cdot u(t)$ | $\tilde{x}_t \cdot z(t)$ |
|---|---|---|---|---|
| 10 | -0.57 | -1.504 | 0.7546 | 1.301 |
| 11 | 0.98 | -0.0449 | -0.0224 | 0.0389 |
| 12 | 1.08 | -0.0358 | -0.0358 | 8E-05 |
| 13 | 3.28 | 2.0733 | 1.0413 | 1.7929 |
| 14 | 2.18 | 0.8824 | -0.4391 | 0.7654 |
| 15 | 1.12 | -0.2685 | 0.2685 | -8E-04 |
| 16 | 0.58 | -0.8994 | 0.4522 | 0.7775 |
| 17 | 1.03 | -0.5403 | -0.2686 | 0.4688 |
| 18 | 2.22 | 0.5588 | 0.5588 | -0.002 |
| sum | | 1.7311 | 1.4213 | 11.31 |

فنجد ما يلي:

$$\hat{\beta} = \frac{2}{19}\sum_{t=0}^{18} \tilde{x}_t \cdot u(t) = \frac{2}{19}(1.4213) = 0.1496 > 0 \quad \& \quad \hat{\gamma} = \frac{2}{19}\sum_{t=0}^{18} \tilde{x}_t \cdot z(t) = \frac{2}{19}(11.31) = 1.1906 > 0$$

و منهما نجد $\hat{\alpha} = 1.19996$ ، وبما أنَّ $\hat{\beta} > 0$ و $\hat{\gamma} > 0$ فإنَّ قيمة $\hat{\phi}$ نجدها باستخدام الجدول (١١ , ١٢) تساوي:

$$\hat{\phi} = -\hat{\theta} = -\mathbf{arctan}\left|\frac{\hat{\gamma}}{\hat{\beta}}\right| = -1.4458$$

وبالتعويض في صيغة المركَّبة $\mathbf{s}(t)$ يكون لدينا:

$$\hat{\mathbf{s}}(t) = \hat{\mathbf{m}} + \hat{\alpha}\ \mathbf{cos}\ (\omega t + \hat{\phi}) = 0.0911 + 1.19998\ \mathbf{cos}\left(\frac{\pi}{3}t - 1.4458\right)$$

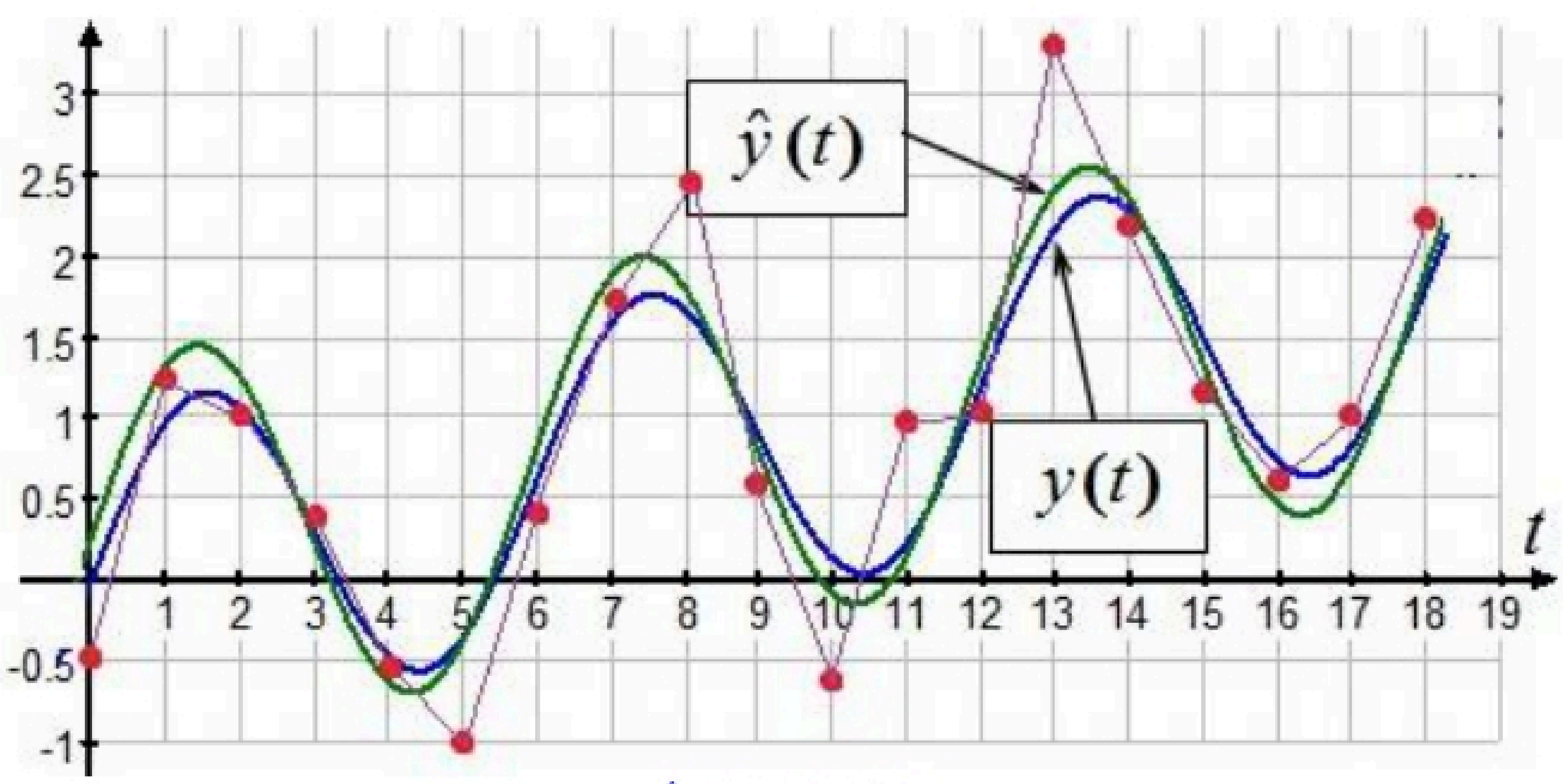
وبالتعويض في صيغة المتسلسلة الزمنية ينتج لدينا أنَّ الطوري العشوائي المقدّر للطوري العشوائي المولّد للمتسلسلة الزمنية هو:

$$\hat{X}_t = \hat{\mathbf{m}}(t)+\hat{\mathbf{s}}(t)+N_t = 0.1161+0.0909\,t+1.19998\,\mathbf{cos}\left(\frac{\pi}{3}t-1.4458\right)+N_t$$

انظر الشكل الآتي (١٧ , ١٢ . ب) لمقارنة بيان الدالَّة $y\,(t)=\mathbf{m}(t)+\mathbf{s}(t)$ مع بيان الدالَّة المقدَّرة لها:

$$\hat{y}\,(t)=\hat{\mathbf{m}}(t)+\hat{\mathbf{s}}(t)=0.1161+0.0909\,t+1.19998\,\mathbf{cos}\left(\frac{\pi}{3}t-1.4458\right) \qquad (*)$$

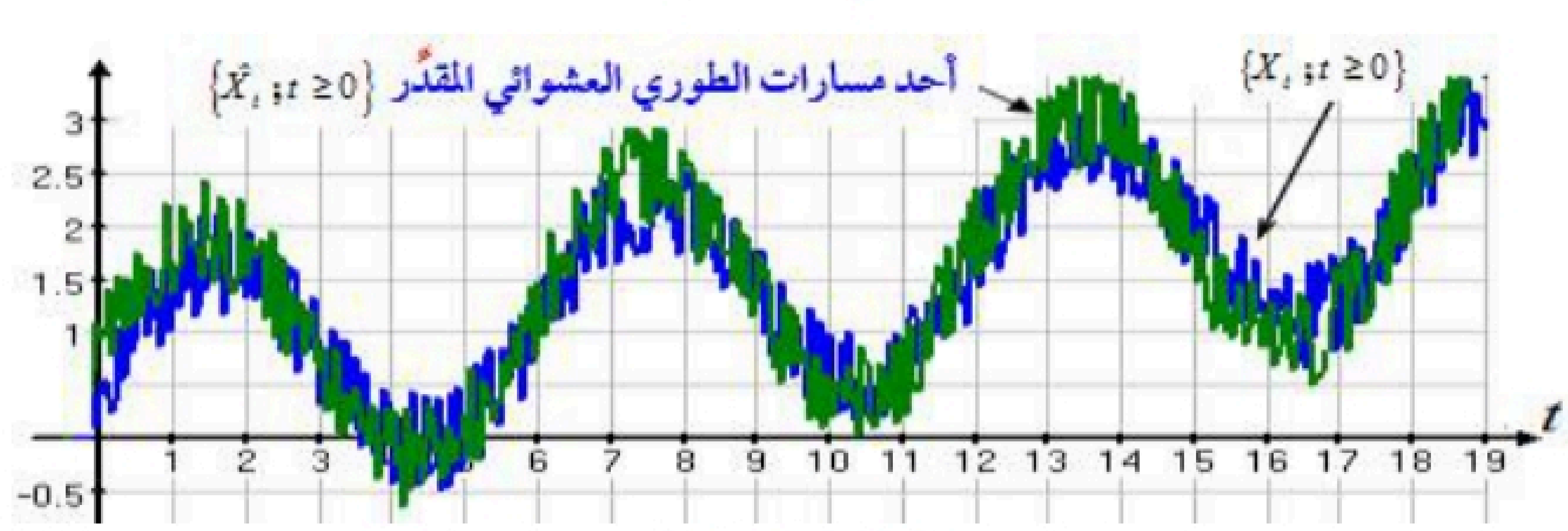
حيث نلاحظ جودة هذه الطريقة في عملية التقدير.




الشكل (١٧ , ١٢ . أ)

بعد إضافة المركِّبة العشوائية إلى العلاقة السابقة (*) نحصل على العرض الآتي لأحد مسارات الطوري العشوائي $\{X_t\,;\,t\ge 0\}$، وأمّا العرض الذي يليه فإنَّه يقدِّم عرضاً نسبياً لمساري الطوريين العشوائيين $\{X_t\,;\,t\ge 0\}$ ومقدِّره $\{\hat{X}_t\,;\,t\ge 0\}$.


الشكل (١٧ , ١٢ . ب)




الشكل (١٧ , ١٢ . ج)

### (٢ , ٤ , ١٢) اختبار جودة التوفيق للمُركّبة الدوريّة المقدَّرة بوساطة طريقة جيب التمام لـ هلبرغ

**Test for Goodness of Fit for the Period Component**

فيما يلي نقدم اختباراً يبيِّن لنا إن كان توفيق تقدير المركّبة الدوريّة الناتجة عن طريقة جيب التمام لـ هلبرغ مقبولاً أم لا. بالطبع يُشترط هنا أيضاً أن تكون أطوال الأدوار للمتسلسلة الزمنية متساوية، أي إنَّ جميع الأدوار لهذه المتسلسلة الزمنية تملك العدد نفسه من المشاهدات (أو البيانات). عندئذ بفرض أنَّ الدالَّة المقدَّرة هي:

$$\hat{y}(t) = \hat{\mathbf{m}}(t) + \hat{\mathbf{s}}(t) \qquad [12,10]$$

وأنَّ البيانات النّاتجة عن هذه الدالَّة هي $\hat{y}_1$ ، $\hat{y}_2$ ،...و $\hat{y}_\tau$ من أجل $t=1$ ، $t=2$ ،...و $t=\tau$ على الترتيب، فإنَّ هذه الطريقة تقوم على تنفيذ الخطوات الآتية:

١- نقوم بوضع البيانات المُقدَّرة $\hat{y}_1$ ، $\hat{y}_2$ ،...و $\hat{y}_\tau$ في جدول بترتيب مصفوفي بحيث يوضع في العمود الأول القيم المُقدَّرة المتتالية في الدور الأول، ومن ثمَّ يوضع في العمود الثاني القيم المقدَّرة المتتالية في الدور الثاني، وهكذا دواليك، فنحصل على العرض الجدولي الآتي:

الجدول (١٣ , ١٢)

| | | الأدوار | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | ... | k |
| المشاهدة ضمن الدور لواحد | 1 | $\hat{y}_{11}$ | $\hat{y}_{12}$ | ... | $\hat{y}_{1k}$ |
| | 2 | $\hat{y}_{21}$ | $\hat{y}_{22}$ | ... | $\hat{y}_{2k}$ |
| | ⋮ | ⋮ | ⋮ | ⋮ ⋮ ⋮ | ⋮ |
| | p | $\hat{y}_{p1}$ | $\hat{y}_{p2}$ | ... | $\hat{y}_{pk}$ |

علماً أنَّ $p$ هو طول الدور، و $k$ هو عدد الأدوار التي لدينا في العيِّنة، وفي حال كان لدينا بيانات غير كافية لشغل العمود الأخير (عدد أقل من $p$ قيمة) فإنَّنا نتجاهل استخدام تلك البيانات الأخيرة.

٢- نقوم باختبار الفرضية الابتدائية الآتية عند مستوى من الأهميّة قدره $\alpha$ :

$\mathbf{H}_o$ : **النموذج المقدَّر للمركبة الدوريّة (الموسميّة) مناسب**

مقابل الفرضية البديلة:

$\mathbf{H}_A$ : **النموذج المقدَّر للمركبة الدوريّة (الموسميّة) غير مناسب**

ومن أجل اتخاذ القرار الخاص بهذا الاختبار نقوم بما يلي:

أ- نحسب المتوسط لكل صف من صفوف الجدول السابق (١٣ , ١٢) بوساطة العلاقة الآتية:

$$\overline{\hat{y}}_i = \frac{1}{k}\sum_{j=1}^{k}\hat{y}_{ij} \qquad ; i = 1, 2, ..., p$$

ب- نقوم بحساب صيغ التباينات الآتية:

$$SSD = \sum_{t=1}^{\tau}(x_t - \hat{y}_t)^2 \qquad \& \qquad SSE = \sum_{i=1}^{p}\sum_{j=1}^{k}\left(\hat{y}_{ij} - \overline{\hat{y}}_i\right)^2 \qquad [12,11]$$

ج- نعرّف إحصاءة $F$ من خلال العلاقة الآتية:

$$F := \frac{(n-p)(SSD - SSE)}{(p-3)SSE} \qquad [12,12]$$

د- نقارن قيمة الإحصاءة $F$ مع القيمة الجدولية $f_{1-\alpha}^{[p-3\,;\,n-p]}$ التي تستخرج من جدول توزيع فيشر.

هـ – اتخاذ القرار: إذا وجدنا أنَّ $F > f_{1-\alpha}^{[p-3\,;\,n-p]}$، فعندئذ تُرفض الفرضية الابتدائية $\mathbf{H}_o$ عند مستوى الأهمّية $\alpha$ ونقبل الفرضية البديلة $\mathbf{H}_A$، وفي حال $F \geq f_{1-\alpha}^{[p-3\,;\,n-p]}$ فإنَّنا نقبل بالفرضية الابتدائية $\mathbf{H}_o$ عند مستوى الأهمّية $\alpha$.

مثال (١٢,٤,٢,١)

بالعودة إلى المثال (١٢,٤,١,١) حيث قمنا بتقدير المركّبة الدوريّة للمتسلسلة الزمنية المولَّدة بالطوري العشوائي $\{X_t\,;\,t \geq 0\}$ المُعطى بالعلاقة الاتية:

$$X_t = 0.1\,t + \sin(2\pi/6)t + N_t \qquad ; t \geq 0$$

ولنبيّن إن كان نموذج جيب التمام لـ هلبرغ المُقدِّر للمركّبة الدوريّة (الموسميّة) في ذلك المثال هو نموذج مناسب أم لا، وذلك عند مستوى من الأهمّية $\alpha = 0.05$ ؟

من أجل ذلك علينا اختبار الفرضية الابتدائية الآتية عند مستوى الأهمّية $\alpha = 0.05$ :

$\mathbf{H}_o$ : **النموذج المقدِّر للمركبة الدوريّة (الموسميّة) مناسب**

مقابل الفرضية البديلة:

$\mathbf{H}_A$ : **النموذج المقدِّر للمركبة الدوريّة (الموسميّة) غير مناسب**

حيث نعلم أنَّه لدينا ثلاثة أدوار في العيِّنة المقدَّمة، وقيمها المقدَّرة وفقاً للدالَّة:

$$\hat{y}\,(t) = \hat{\mathbf{m}}\,(t) + \hat{\mathbf{s}}\,(t) = 0.1161 + 0.0909\,t + 1.19998\,\cos\left(\frac{\pi}{3}t - 1.4458\right)$$

يقدِّمها لنا الجدول الآتي:

الجدول (١٢,١٤)

| | | Periods الأدوار | | | sum | $\bar{\hat{y}}_i$ |
|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | | |
| $\bar{\hat{y}}_i$ القيم المقدَّرة المتتالية في كل دور | 1 | 0.266 | 0.81 | 1.354 | 2.43 | 0.81 |
| | 2 | 1.313 | 1.858 | 2.403 | 5.57 | 1.858 |
| | 3 | 1.254 | 1.801 | 2.347 | 5.4 | 1.801 |
| | 4 | 0.24 | 0.787 | 1.334 | 2.36 | 0.787 |
| | 5 | -0.626 | -0.08 | 0.466 | -0.2 | -0.08 |
| | 6 | -0.386 | 0.158 | 0.703 | 0.48 | 0.158 |

وبتنفيذ الحسابات المطلوبة لـ $SSD$ و $SSE$ نجد باستخدام العلاقتين [12,11] أنّ:

$$SSD = 3.6846 \qquad \& \qquad SSE = 3.5709$$

ومن ثمَّ يكون لدينا بحسب العلاقة [12,12]:

$$F = \frac{(18-6)(3.6846-3.5709)}{(6-3)\;3.5709} = \frac{1.3644}{6.5709} = 0.208$$

ومن جدول توزيع فيشر نجد أنَّ $f_{0.95}^{[3;12]} = f_{1-\alpha}^{[p-3\,;\,n-p]} = 3.49$، وهكذا يكون لدينا:

$$f_{0.95}^{[3;12]} = 3.49 > F = 0.208$$

ولذلك علينا القبول بالفرضية الابتدائية عند مستوى الأهمّية $\alpha = 0.05$ .

في الواقع إنَّ النموذج الذي قمنا بتقديره للمركبة الدوريّة للمتسلسلة الزمنية في هذا المثال هو نموذج مناسب جداً لتلك المُركَّبة ويتّضح ذلك بجلاء من خلال أخذ قيمة كبيرة لـ $\alpha$ كأن نأخذ على سبيل المثال $\alpha = 0.10$ ، فنجد رغم ذلك أنَّ قيمة الإحصاءة $F$ أصغر من القيمة الجدولية حيث لدينا:

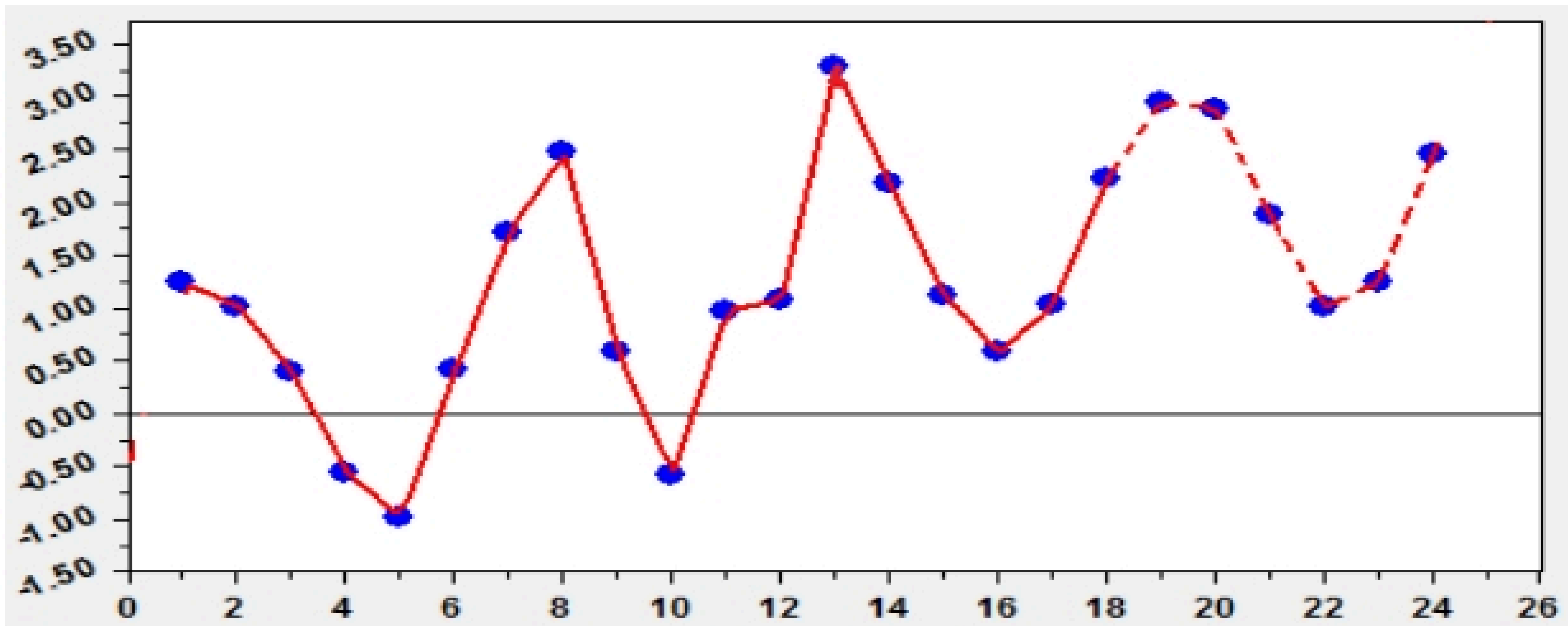
$$f_{1-\alpha}^{[p-3\,;\,n-p]} = f_{0.90}^{[3;12]} = 2.61$$

الآن وبعد التّحقّق من أنَّ النموذج الذي توصلنا إليه مناسب، فإنَّه يمكننا توليد تنبؤات من هذا النموذج (وكما ذكرنا سابقاً على ألا يكون لأكثر من طول دور واحد)، فنجد على سبيل المثال قيم التنبؤات حتى طول دور كامل هي:

الجدول (١٥ ,١٢)

| Lag | 19 | 20 | 21 | 22 | 23 | 24 |
|---|---|---|---|---|---|---|
| $x_t$ | 2.95 | 2.89 | 1.88 | 1.01 | 1.25 | 2.45 |

ولعرضها البياني الشكل الآتي:




الشكل (١٨ ,١٢ . أ)

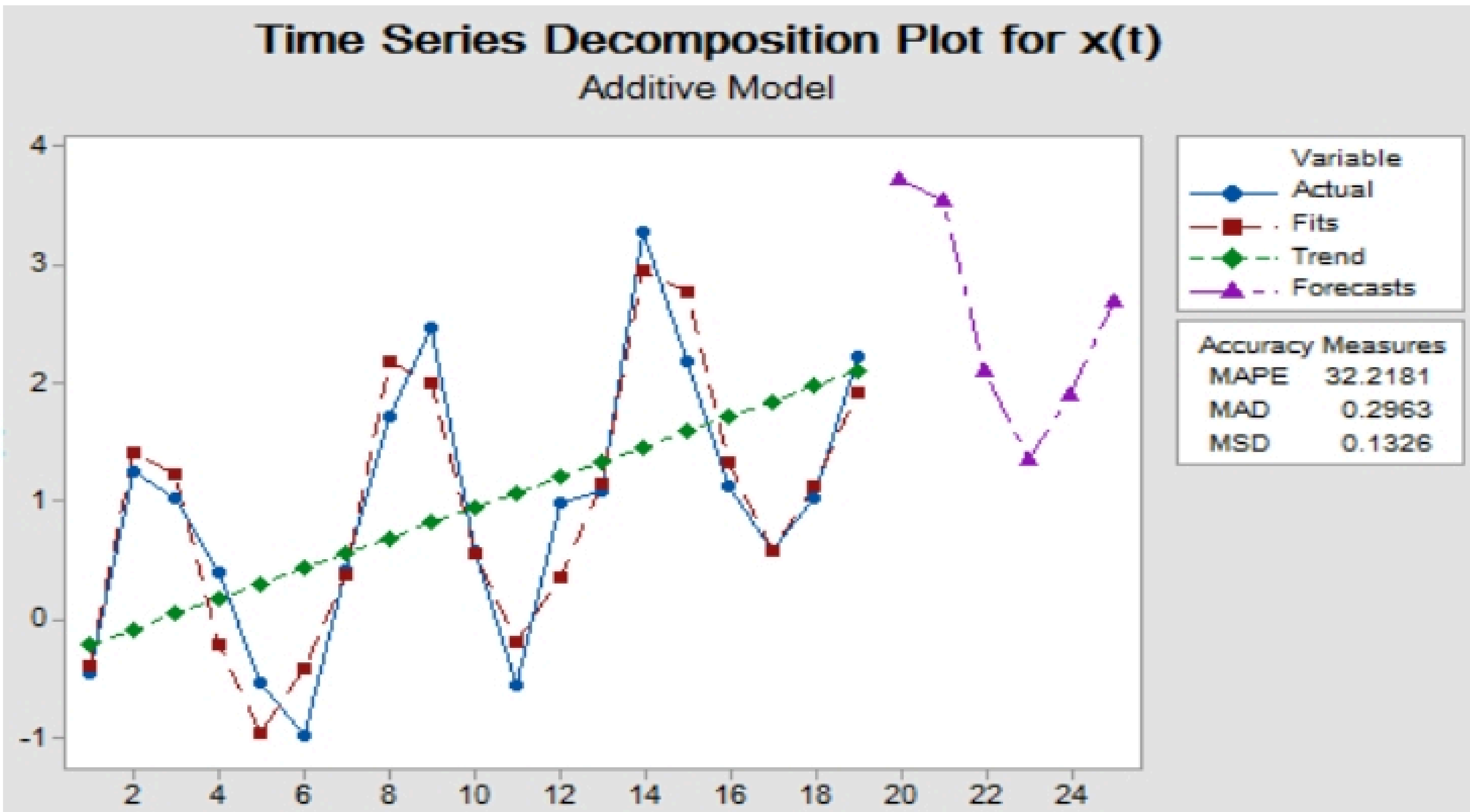
فنلاحظ أنَّ هذه التنبؤات مقبولة إلى حدٍّ ما بالمقارنة مع سلوك هذه العيِّنة من مسار المتسلسلة الزمنية التي هي قيد الدراسة.

العرض الآتي (الشكل (١٨ ,١٢ . ب)) يقدِّم لنا التوفيق (القيم المقدَّرة) لبيانات هذه العيِّنة مع التنبؤات لدور كامل باستخدام برنامج Minitab وذلك وفقاً لطريقة التجزئة (نموذج الجمع)، حيث يظهر لنا الاتجاه العام في هذه اللوحة أيضاً، وكذلك نلاحظ أنَّ قيم التنبؤات التي قدَّمها هذا البرنامج تُعدّ أكثر منطقية من تلك التي قدمناها وفقاً لطريقة جيب التمام لـ هلبرغ.

**Forecasts**

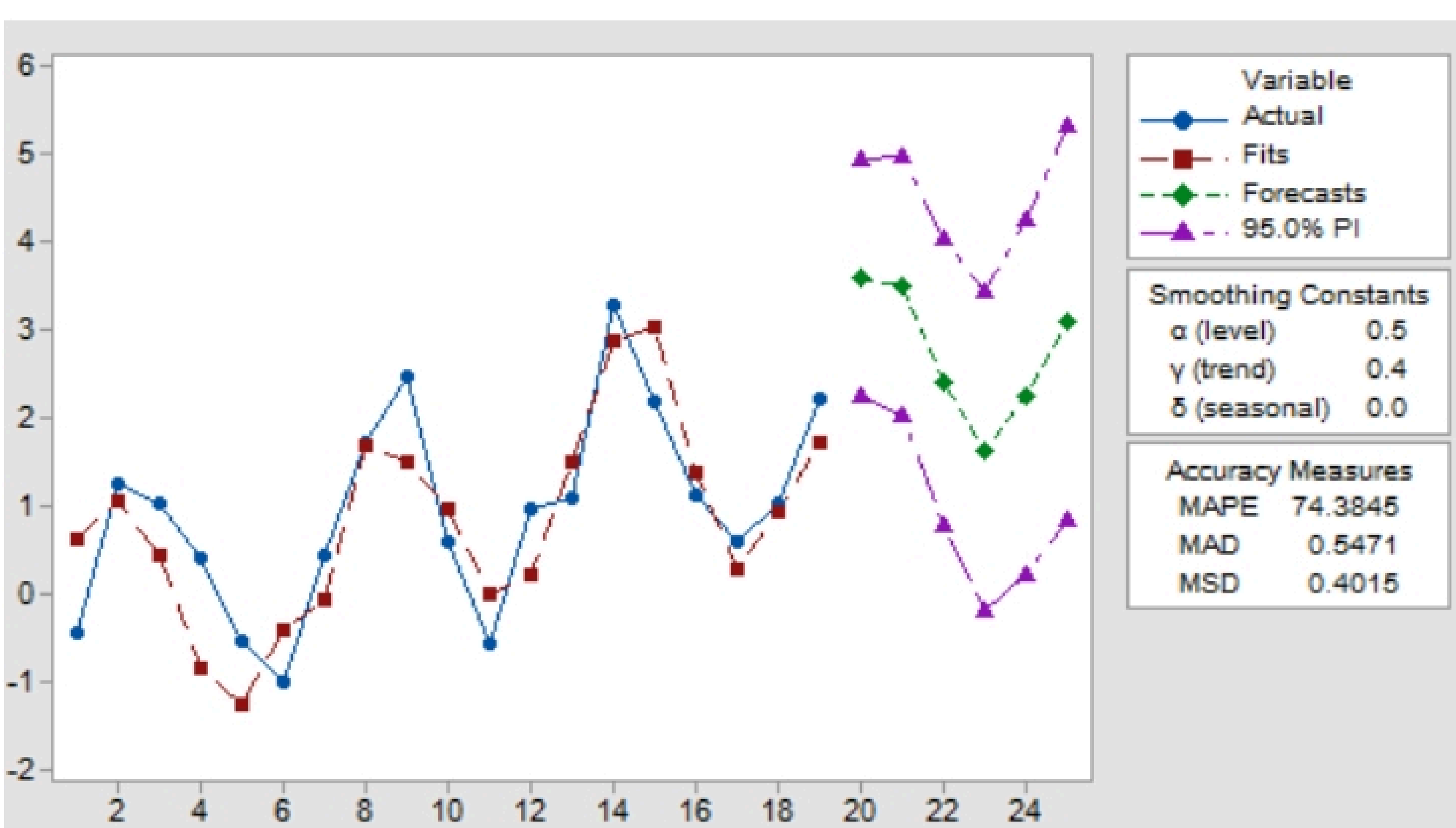
| Period | Forecast |
|---|---|
| 20 | 3.72693 |
| 21 | 3.54932 |
| 22 | 2.09963 |
| 23 | 1.34243 |
| 24 | 1.89274 |
| 25 | 2.69763 |

علماً أنَّ **Period** تعني هنا ترتيب القيمة أو الفجوة الزمنية التي أُخِذَ عندها التنبؤ حيث لدينا 19 قيمة سابقة (عدد البيانات الأصل لهذه العيِّنة).




الشكل (١٨, ١٢.ب)

أمّا لإظهار فترة الثقة التي نوّهنا عنها سابقاً فإنّنا سنستخدم طريقة وينتر Winters' Method (إحدى طرائق التنعيم للمتسلسلات الزمنية) في برنامج Minitab لتوليد التنبؤات، وذلك لأنَّ هذه الطريقة من الطرائق الجيدة والمناسبة لدراسة مثل هذا النوع من المتسلسلات الزمنية، وكذلك تقدِّم هذه الطريقة لنا فترة الثقة للتنبؤات. بعد تحميل طول الدور والتعديل على الأوزان $\alpha$ ، $\gamma$ و $\delta$ بحيث نحصل على أصغر خطأ معيّاري للتقدير MSD، فإنّنا سنحصل على العرض الآتي الذي يتضمَّن 95% فترة ثقة للتنبؤات المولَّدة.




الشكل (١٨, ١٢.ج)

إنَّ استخدام هذه الطريقة يتطلب خبرة في كيفية تغيير قيم الأوزان $\alpha$، $\gamma$ و $\delta$ الخاصّة بهذه الطريقة، وسوف نقدِّم شرحاً موجزاً حول هذه الطريقة في نهاية هذا الفصل.

## (١٢,٥) تنعيم المتسلسلات الزمنية

**Smoothing of Time Series**

قد تصادفنا حالات يكون فيها الاهتمام منصباً على المنحى الذي تتخذه المتسلسلة الزمنية وليس على التذبذبات (سواء كانت دوريّة أو موسميّة) الحاصلة في مسار هذه المتسلسلة الزمنية. في مثل هذه الحالات يقوم المرء بمحاولة التخلص من هذه التذبذبات من خلال ما يُعرف باسم تنعيم المتسلسلة الزمنية التي تؤدي بالنتيجة إلى متسلسلة أخرى ذات تذبذبات بمطالات صغيرة بالمقارنة مع تذبذبات المتسلسلة الزمنية الأصل، ولكن لكي نشرح هذه الطريقة سنحتاج إلى تقديم بعض المفاهيم التي سندرجها من خلال التعاريف الآتية.

### (١٢,٥,١) تعريف (مُعامل الخشونة Roughness Coefficient)

لتكن لدينا $x_1$، $x_2$،...و $x_\tau$ بيانات متسلسلة زمنية. عندئذ يُعرَّف معامل الخشونة لهذه المتسلسلة الزمنية بوساطة العلاقة الآتية:

$$\mathbf{RC} := \sum_{t=2}^{\tau}\left(x_t - x_{t-1}\right)^2 \Big/ \sum_{t=2}^{\tau}\left(x_t - \bar{x}\right)^2 \qquad [12,13]$$

لاحظ أنَّ هذه العلاقة هي نسبة متوسط مربَّع انحرافات القيم المتتالية للبيانات إلى تباين البيانات (حول متوسطها)، ومن ثَمَّ قيمة هذا المُعامل يزداد كبراً كلّما كَبُر الفرق بين القيم المتتالية وصَغُر بين القيم ومتوسطها.

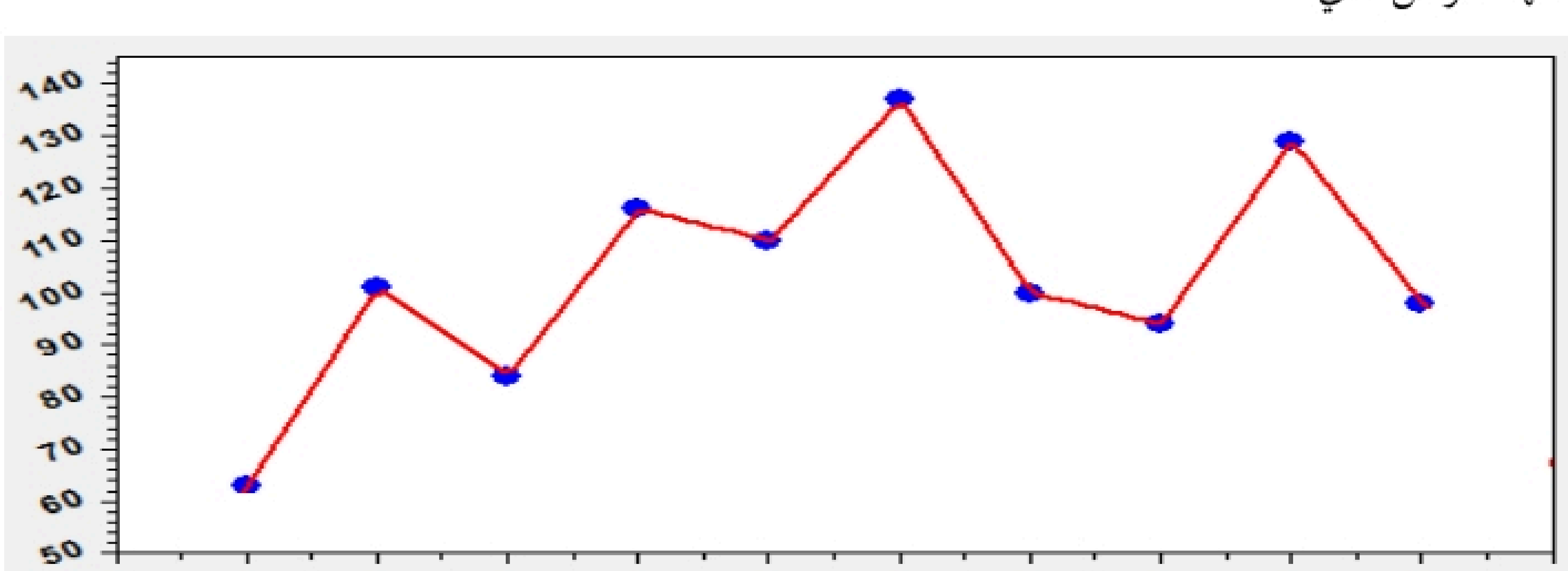
### (١٢,٥,١,١) مثال

لتكن لدينا البيانات الآتية لعينة مأخوذة من مسار متسلسلة زمنية:

الجدول (١٢,١٦)

| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 63 | 101 | 84 | 116 | 110 | 137 | 100 | 94 | 129 | 98 |

والتي لبصمتها العرض الآتي:




شكل (١٢,١٩).

فنجد أنَّ متوسط البيانات المُعطاة يساوي $\overline{x} = 103.2$، ومن ثَمَّ بحسب العلاقة [12,13] تكون قيمة مُعامل الخشونة لهذه المتسلسلة الزمنية يساوي:

$$\mathbf{RC} = 7113 \times \left(2513.56\right)^{-1} = 2.83$$

### (١٢,٥,٢) التنعيم باستخدام المتوسطات المتحركة Smoothing by Moving Averages

في الواقع يمكن تنعيم مسار متسلسلة زمنية تعاني من خشونة كبيرة بوساطة ما يسمَّى بالمتوسطات المتحركة ذات الطول $k$ التي تقدِّمها لنا الفقرة الآتية.

#### (١٢,٥,٢,١) تعريف المتوسط المتحرك بطول $k$

لتكن لدينا $x_1$، $x_2$،...و $x_\tau$ بيانات مُعطاة، و $k \in \mathbb{N}$ عدد مُعطى. عندئذ يُعرَّف المتوسط المتحرك بطول $k$ لهذه البيانات بوساطة العلاقة الآتية:

$$\bar{x}_{k,t} := \frac{x_t + x_{t+1} + \dots + x_{t-\tau+1}}{k} \quad ; k = 1, 2, 3, \dots, \tau - (k-1) \qquad [12,14]$$

#### (١٢,٥,٢,٢) ملاحظات

١- إنَّ البيانات الناتجة عن قيم المتوسطات المتحركة بطول $k$ لمتسلسلة زمنية تُعطينا بدورها متسلسلة زمنية جديدة ولكن بمُعَامل خشونة أقل (على الغالب) من معامل الخشونة للمتسلسلة الزمنية الأصل، ولهذا السبب ستكون المتسلسلة الزمنية المولَّدة بمتوسطات متحركة أكثر نعومةً من المتسلسلة الزمنية الأصل على الغالب، ولذلك تستخدم طريقة المتوسطات المتحركة كإحدى الطرائق لتنعيم المتسلسلات الزمنية، وفي حال اللجوء إلى هذه الطريقة يُؤخذ عادة طول الدور (أو عدد المواسم في العام) كطول للمتوسط المتحرك الذي سيستخدم.

٢- إذا كان طول المتوسط المتحرك المستخدم $k$، فعندئذ سينقص عدد البيانات الناتجة عن استخدام هذا المتوسط المتحرك بمقدار $\tau-(k-1)$، ولتوضيح الملاحظتين السابقتين سنأخذ المثال الآتي.

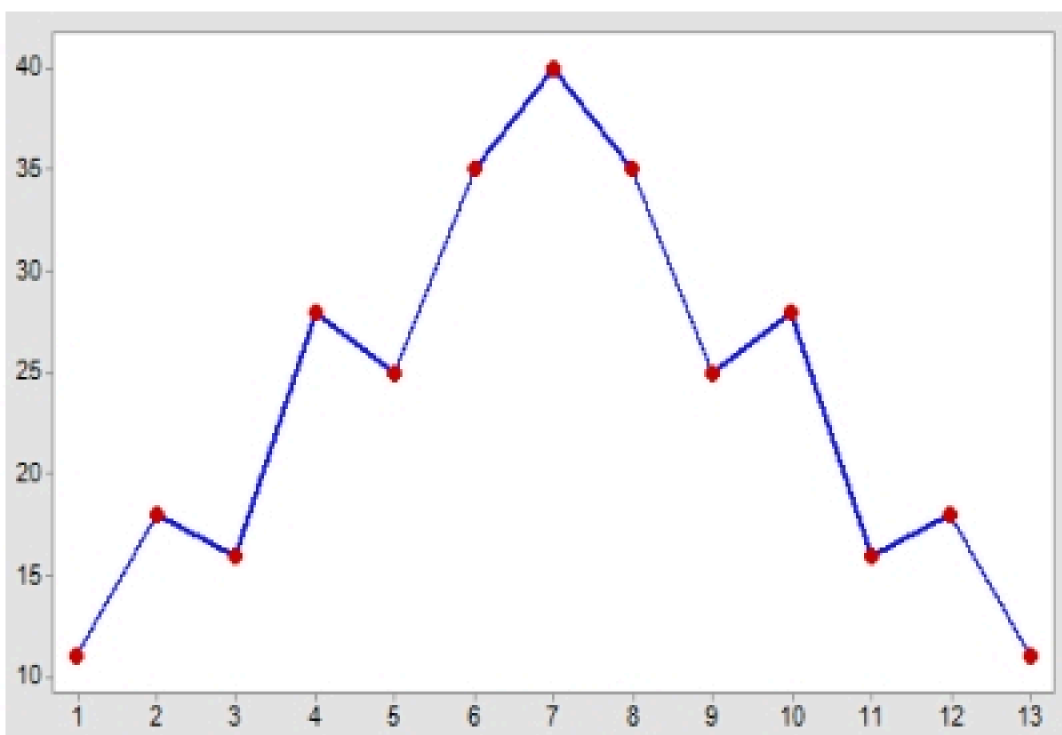
#### (١٢,٥,٢,٣) مثال

لتكن لدينا عيِّنة مأخوذة من مسار متسلسلة زمنية مُعطاة كما في الجدول الآتي:

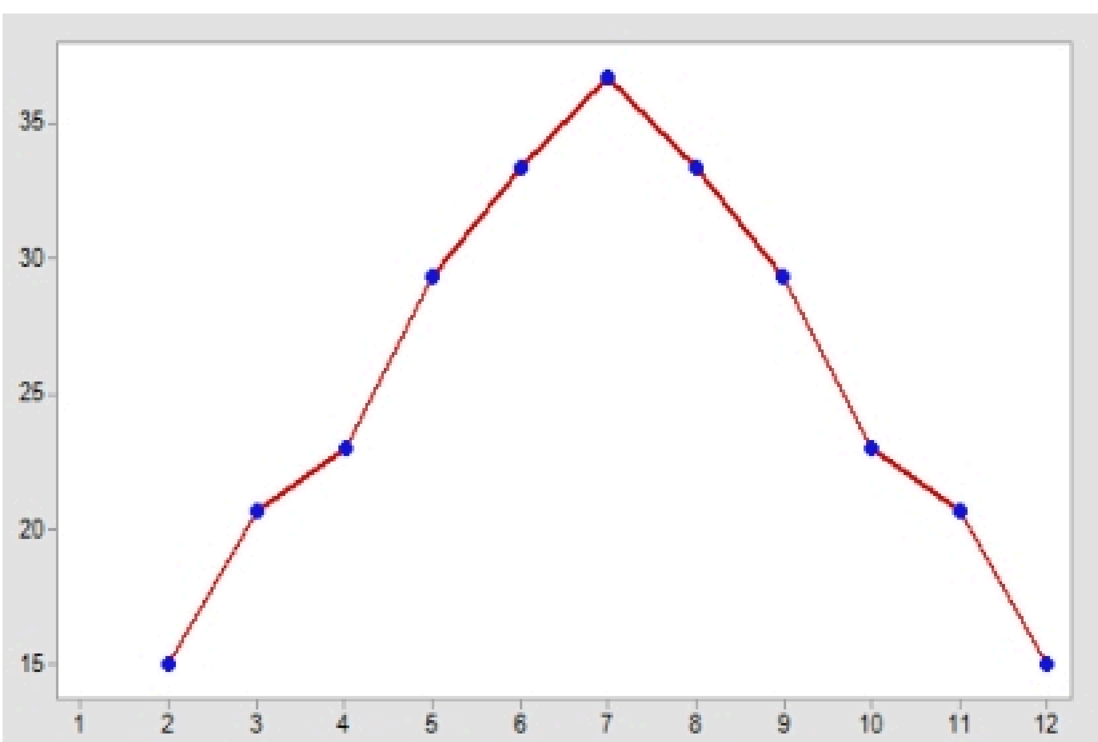
الجدول (١٧ ,١٢. أ)

| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 11 | 18 | 16 | 28 | 25 | 35 | 40 | 35 | 25 | 28 | 16 | 18 | 11 |

والتي لبصمتها الشكل (٢٠ ,١٢. أ).



الشكل (٢٠ ,١٢. أ)



الشكل (٢٠ ,١٢. ب)

ومن أجل حساب معامل الخشونة لهذه المتسلسلة الزمنية لنبني الجدول الآتي:

الجدول (١٧ , ١٢. ب)

| $t$ | $x_t$ | $(x_t - x_{t-1})^2$ | $(x_t - \bar{x})^2$ |
|---|---|---|---|
| 1 | 11 | | |
| 2 | 18 | 49 | 30.6916 |
| 3 | 16 | 4 | 56.8516 |
| 4 | 28 | 144 | 19.8916 |
| 5 | 25 | 9 | 2.1316 |
| 6 | 35 | 100 | 131.3316 |
| 7 | 40 | 25 | 270.9316 |
| 8 | 35 | 25 | 131.3316 |
| 9 | 25 | 100 | 2.1316 |
| 10 | 28 | 9 | 19.8916 |
| 11 | 16 | 144 | 56.8516 |
| 12 | 18 | 4 | 30.6916 |
| 13 | 11 | 49 | 157.2516 |
| sum | 306 | 662 | 909.9792 |

فنجد أنَّ قيمة متوسط البيانات الأصل $\overline{x} = 23.53846$ ، ومن ثمَّ بحسب العلاقة [12,13] يكون لدينا $\mathbf{RC} = \dfrac{662}{909.979} = 0.727$ ، ولنقم بتوليد متسلسلة زمنية من المتسلسلة الزمنية السابقة باستخدام متوسط متحرك بطول 3 فنجد بياناتها كما في الجدول الآتي:

الجدول (١٨ , ١٢)

| $t$ | $x_t$ | $(x_t - x_{t-1})^2$ | $(x_t - \bar{x})^2$ |
|---|---|---|---|
| 1 | ---------- | ---------- | ---------- |
| 2 | 15 | ---------- | ---------- |
| 3 | 20.66667 | 32.11111 | 22.30988 |
| 4 | 23 | 5.444444 | 5.7121 |
| 5 | 29.33333 | 40.11111 | 15.54988 |
| 6 | 33.33333 | 16 | 63.09654 |
| 7 | 36.66667 | 11.11111 | 127.1632 |
| 8 | 33.33333 | 11.11111 | 63.09654 |
| 9 | 29.33333 | 16 | 15.54988 |
| 10 | 23 | 40.11111 | 5.7121 |
| 11 | 20.66667 | 5.444444 | 22.30988 |
| 12 | 15 | 32.11111 | 107.9521 |
| 13 | ---------- | ---------- | ---------- |
| sum | 279.3333 | 209.5556 | 448.4521 |

فنجد أنَّ قيمة متوسط البيانات الأصل $\overline{x} = 25.394$، وبحسب العلاقة [12,13] يكون معامل الخشونة لهذه المتسلسلة الزمنية يساوي $\mathbf{RC} = 0.467$ ، والشكل السابق (٢٠ , ١٢. ب) يعرض لنا بصمة المتسلسلة الزمنية المولَّدة بالمتوسط المتحرك ذي الطول 3 (لاحظ النعومة التي طرأت على مسارها).

لاحظ أنَّه لتبسيط العمليّات الحسابية فإنَّه عند تشكيل المتسلسلة الزمنية المولَّدة بمتوسط متحرك نقوم بترتب قيمها بحيث يكون عدد الفراغات (الناتجة عن النقص في عدد المشاهدات) في الأعلى يساوي عدد الفراغات التي في الأسفل إذا كان طول المتوسط المتحرك $k$ عدداً فردياً، وأمّا إذا كان $k$ عدداً زوجياً فعندئذ نجعل الفراغات في الأعلى أقل من الفراغات التي في الأسفل بمقدار فجوة زمنية واحدة. بالطبع هناك الكثير من الدراسات التي تتعلّق بالمركّبات الدوريّة للمتسلسلات الزمنية ولكنّنا سنكتفي بهذا القدر من الدراسة، ويمكن لمن يودّ الاطلاع على المزيد الرجوع إلى المراجع ذات الصلة ومنها في فهرس المراجع في نهاية هذا الكتاب.

## (١٢,٦) دراسة التغيرات الموسمية لمتسلسلة زمنية

**Study of the Seasonal Changes of a Time Series**

لقد ذكرنا سابقاً أنَّ التغيّرات الموسمّية (التي تمثِّلها المركِّبة الموسميّة) لمتسلسلة زمنية هي تغيّرات متشابهة ذات تكرار منتظم على مدى فترة من الزمن (تُدعى طول الفترة -أو الدورة-الموسميّة)، وهنا نودّ أن نشير إلى أنَّ كلمة موسميّة قد تأخذ دلالات مختلفة للوصف السابق، فعلى سبيل المثال لا الحصر:

١- من الممكن أن تحدث التغيّرات الموسميّة كل يوم مثل قيم درجات الحرارة منتصف النهار إذا كانت البيانات المقدَّمة هي قياس درجة الحرارة خلال كل ساعة.

٢- من الممكن أن تحدث التغيّرات الموسميّة كل أسبوع مثل عدد السيارات التي تسير على الطريق الدائري الشرقي في الرياض يوم الجمعة من الساعة السابعة وحتى التاسعة صباحاً إذا كانت البيانات المقدَّمة هو عدد السيارات المارة على هذا الطريق في هذه الفترة الزمنية من كل يوم.

٣- من الممكن أن تحدث التغيّرات الموسميّة في أوقات مُحدَّدة من العام مثل مبيعات متجر ملابس إذا كانت البيانات المقدَّمة هي قيمة المبيعات كل أسبوع، فحينئذ يلاحظ أنَّ الأسبوع الذي يسبق كل عيد (عيد الفطر والأضحى) يبدي تغيّراً موسمياً (بطول فترة موسميّة غير ثابت).

في الحقيقة لوحظ أنَّ المتسلسلات الزمنية في القطاعين الاقتصادي والاجتماعي يكون لها تغيُّرات موسميّة بمواسم شهرية (إذا كانت البيانات المقدَّمة مأخوذة كقياسات يومية) أو ربع سنوية (إذا كانت البيانات المقدَّمة مأخوذة كقياسات يومية أو أسبوعية)، ويعود ذلك إلى أسباب تؤثّر في قيم هذه القياسات من أهمّها العادات والتقاليد الاجتماعية والطقس أيضاً. أمّا المتسلسلات الزمنية في القطاعين الطبي والنفسي فيكون لها تغيُّرات موسميّة بمواسم قد تبدأ بـ ثوان(مثل البيانات التي تستخدم لرسم مخطط الرسم الكهربي للقلب) ومن الممكن أن تمتدّ إلى أشهر وسنوات مثل البيانات التي تمثّل تطور حالة مريض نفسي.

في الواقع يصعب (لا بل من الصعب جداً) تحديد تأثير التغيّرات الموسميّة دون الانتباه إلى الاتجاه العام للمتسلسلة الزمنية، ولذلك يُنصح بإزالة الاتجاه العام للمتسلسلة الزمنية أولاً، ومن ثمَّ البحث في دراسة الأثر الموسمي.

إنَّ دراسة التغيّرات الموسميّة للمتسلسلات الزمنية زاخر بالطرائق المستخدمة لأجله، ولكلّ طريقة هدف قد يختلف عن هدف الطريقة الأخرى، فمنها يستخدم للتخلص من الأثر الموسمي (التخلص من المركِّبة الموسميّة) أو التقليل من شأنه، وبعضها الآخر يقوم على تقدير الأثر الموسمي. هذا من جانب، ومن جانب آخر هناك تباين كبير في صعوبة دراسة الأثر الموسمي وفقاً للهدف والنموذج الذي يخضع للدراسة، فمنها يتطلب التعامل مع منظومة معادلات رياضياتية قد لا تكون بسيطة كما هو الحال في نماذج الانحدار الذاتي والمتوسطات المتحركة التكاملية ARIMA، ومنها ما هو بسيط نسبياً كتلك الطرائق التي تدرس المركِّبة الموسميّة بأسلوب تقليدي، ومن هذه الطرائق:

أ- طريقة المتوسطات المتحركة.

ب- طريقة التعديل الذاتي.

ج- طريقة تحديد التغيّرات الموسميّة.

د- طريقة إيجاد التقلبات الموسميّة بدلالة الوحدات المطلقة.

فيما يلي سنكتفي بشرح طريقة النسبة إلى متوسط متحرك في دراسة المركِّبة الموسميّة ومن يودّ الاطلاع على المزيد يمكنه الرجوع إلى المراجع ذات الصلة.

### (١, ٦, ١٢) طريقة النسبة إلى متوسط متحرك Ratio to Moving-Average Method

لقد لاحظنا أنَّ طريقة المتوسطات المتحركة مفيدة في تنعيم المتسلسلة الزمنية والتخلّص من الاهتزازات المكوِّنة للمركَّبة الدوريّة عندما يكون اهتمامنا منصباً على الاتجاه العام للمتسلسلة الزمنية، ولكنَّ هذه الطريقة مفيدة في مجال تحليل التغيُّرات الموسميّة أيضاً، وتعتمد في تحليلها للتغيُّرات الموسميَّة على الخطوات الآتية:

١- جعل مسار المتسلسلة الزمنية أملساً بالقدر المستطاع من خلال أخذ متوسط متحرك مناسب، ويؤخذ في هذه الحالة طول المتوسط المتحرك يساوي عدد الفصول في العام، مع العلم أنَّ كلمة فصل هنا قد لا تعني بالضرورة ثلاثة أشهر كما هو دارج، وإنّما يعني فترة من السنة يتفق عليها من أجل أخذ القياسات (القيم المشاهدة).

٢- إيجاد حاصل ضرب المركَّبة الموسميّة مع المركَّبة العشوائية وذلك من خلال قسمة البيانات الأصلية للمتسلسلة الزمنية على المتوسط المتحرك المقابل لتلك القيم، ومن ثمَّ ضرب الناتج بـ 100.

٣- عزل المركَّبة الموسميّة عن المركَّبة العشوائية، وذلك بالتعديل على السنوات من خلال جعل القيم الناتجة في الخطوة السابقة لكلّ موسم بعضها فوق بعض لكل السنوات.

٤- حساب المتوسط لقيم الموسم (أو فصل) $k$ للبيانات الناتجة في البند الثاني، حيث سنرمز لهذا المتوسط بـ $\overline{x}_k$ و $k$ تأخذ قيمها بعدد المواسم المُتاحة للدراسة.

٥- حساب ما يُعرف باسم **المجموع الأصلي للنسب المئوية**، وهو عدد المواسم (أو الفصول) مضروب بـ 100، وسنرمز له بـ $\alpha$.

٦- حساب المجموع لكلّ هذه المتوسطات $\overline{x}_k$ (المحسوبة في الخطوة الرابعة) الأخيرة وسنفترضه يساوي $\beta$.

٧- تقدير المركَّبة الموسميّة ذات الرقم $i$ (وسنرمز لها بـ $\hat{s}_i$) من خلال ضرب قيمة المتوسط للموسم (أو الفصل) في القيمة $\gamma = \alpha / \beta$.

### (١, ١, ٦, ١٢) مثال

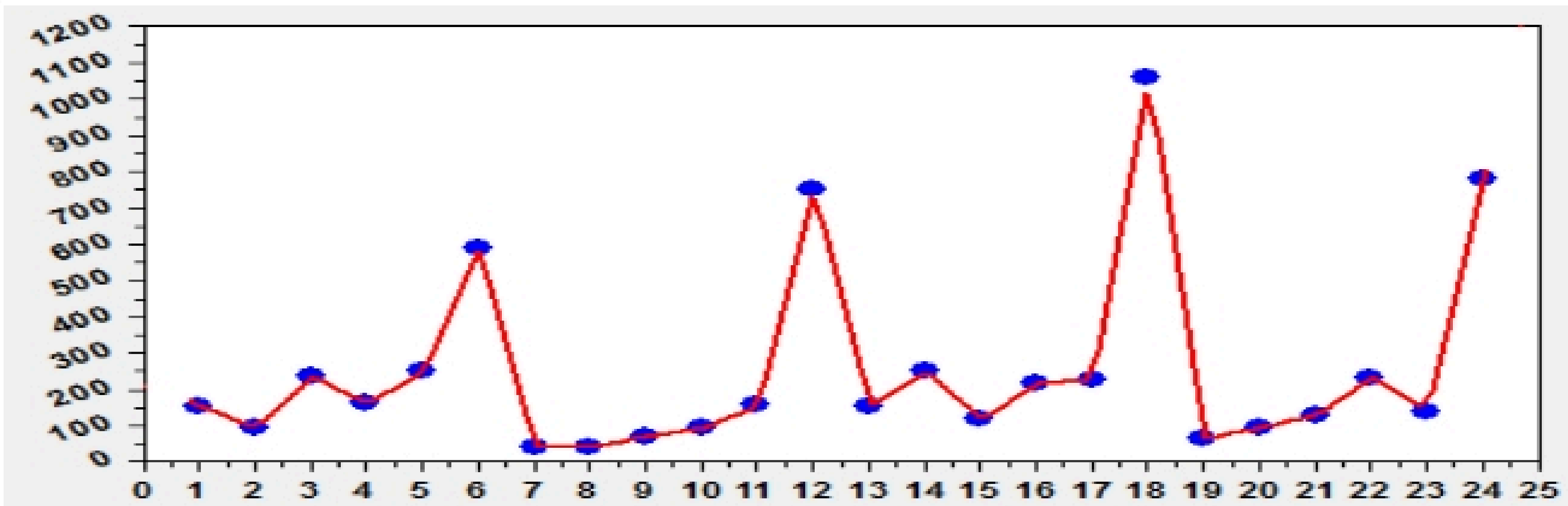
لتكن لدينا البيانات الآتية لعينة مأخوذة من مسار متسلسلة زمنية لمصاريف شركة تجارية على العمال في نهاية كل شهرين (مقدَّرةً بالـ 1000 وحدة نقدية)، أي إنَّه لدينا ستة فصول أو مواسم (طول كل فصل يساوي شهرين) في هذه المتسلسلة الزمنية.

الجدول (١٩, ١٢. أ)

| العام | $t$ | الموسم Season | المصاريف Outlay |
|---|---|---|---|
| 2010 | 1 | 1 | 155 |
| | 2 | 2 | 95 |
| | 3 | 3 | 236 |
| | 4 | 4 | 163 |
| | 5 | 5 | 250 |
| | 6 | 6 | 588 |
| 2011 | 7 | 1 | 41 |
| | 8 | 2 | 41 |
| | 9 | 3 | 71 |
| | 10 | 4 | 94 |
| | 11 | 5 | 157 |
| | 12 | 6 | 751 |

| العام | $t$ | الموسم Season | المصاريف Outlay |
|---|---|---|---|
| 2012 | 13 | 1 | 155 |
| | 14 | 2 | 250 |
| | 15 | 3 | 119 |
| | 16 | 4 | 216 |
| | 17 | 5 | 229 |
| | 18 | 6 | 1061 |
| 2013 | 19 | 1 | 66 |
| | 20 | 2 | 95 |
| | 21 | 3 | 131 |
| | 22 | 4 | 232 |
| | 23 | 5 | 142 |
| | 24 | 6 | 783 |

والتي لبصمتها العرض الآتي:



الشكل (٢٠,١٢.ب)

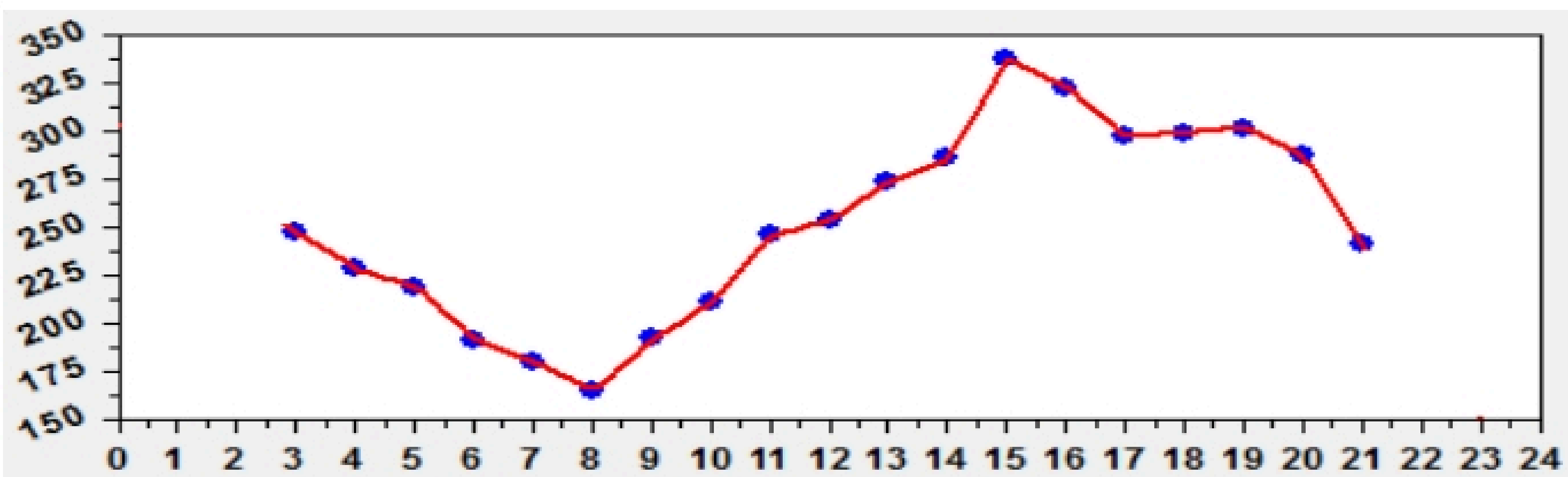
ولنقم بتقدير المركِّبات الموسميّة لهذه المتسلسلة الزمنية وفقاً لطريقة النسبة إلى متوسط متحرك.

من أجل تحقيق البند الأول سنقوم بتنعيم المتسلسلة الزمنية التي لدينا باستخدام متوسط متحرك بطول يساوي 6 (لأنَّه لدينا ست مشاهدات في كل عام) فنحصل على المتسلسلة الزمنية المنعَّمة الآتية:

الجدول (١٩,١٢.ب)

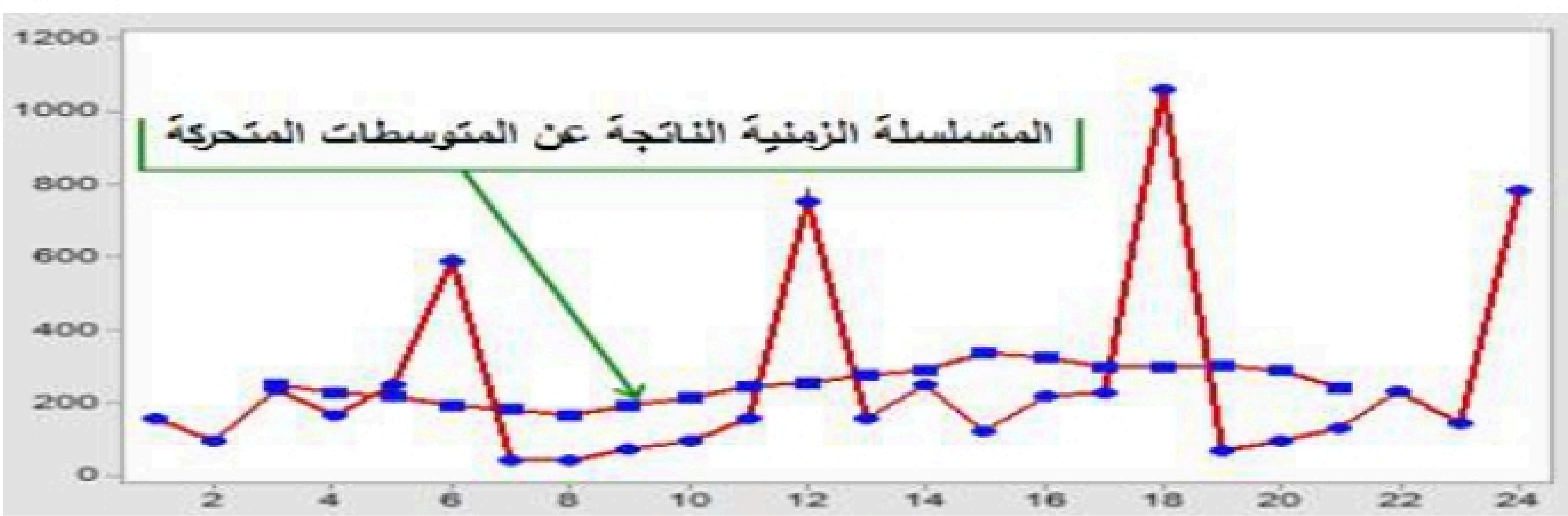
| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $\bar{x}_{6,t}$ | ------- | ------- | 247.83 | 228.83 | 219.83 | 192.33 | 180.83 |
| $t$ | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| $\bar{x}_{6,t}$ | 165.33 | 192.5 | 211.5 | 246.33 | 254.33 | 274.67 | 286.67 |
| $t$ | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| $\bar{x}_{6,t}$ | 338.33 | 323.5 | 297.67 | 299.67 | 302.33 | 287.83 | 241.5 |

والتي لبصمتها الشكل الآتي:



الشكل (٢٠,١٢.ج)

ووضعها النسبي مع المتسلسلة الزمنية الأصل يقدِّمها الشكل الآتي:




الشكل (٢٠,١٢.د)

لاحظ كيف أصبح مسار المتسلسلة الزمنية الناتجة عن المتوسطات المتحركة ناعماً جداً بالمقارنة مع مسار المتسلسلة الزمنية الأصل.

الآن لتطبيق الخطوة الثانية من هذه الطريقة نقوم ببناء الجدول الآتي حيث نجد فيه نواتج حاصل ضرب المركّبة الموسميّة مع المركّبة العشوائية التي تنتج من خلال قسمة البيانات الأصلية للمتسلسلة الزمنية على المتوسط المتحرك المقابل لتلك القيم، ومن ثمّ ضرب الناتج بـ 100.

الجدول (١٩, ١٢. ج)

| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $(x_t/\bar{x}_{6,t})\times 100$ | ------- | ------- | 95.2 | 71.23 | 114 | 305.7 | 22.67 |
| $t$ | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| $(x_t/\bar{x}_{6,t})\times 100$ | 24.8 | 36.88 | 44.44 | 63.7 | 295.3 | 56.4 | 87.21 |
| $t$ | 15 | 16 | 17 | 18 | 19 | 20 | 21 |
| $(x_t/\bar{x}_{6,t})\times 100$ | 35.17 | 66.8 | 76.93 | 354.1 | 21.8 | 33.01 | 54.2 |

ولكي نعزل المركّبة الموسميّة عن المركّبة العشوائية نقوم بالتعديل على السنوات من خلال جعل القيم الناتجة في الخطوة السابقة لكلّ موسم بعضها فوق بعض لكلّ السنوات، وناتج هذه العمليّات يبيّنها الجدول الآتي:

الجدول (١٩, ١٢. د)

| Year | الموسم الأول | الموسم الثاني | الموسم الثالث | الموسم الرابع | الموسم الخامس | الموسم السادس | |
|---|---|---|---|---|---|---|---|
| 2010 | ----- | ----- | 95.23 | 71.23 | 113.7 | 305.7 | |
| 2011 | 22.67 | 24.8 | 36.88 | 44.44 | 63.73 | 295.3 | |
| 2012 | 56.43 | 87.21 | 35.17 | 66.77 | 76.93 | 354.1 | |
| 2013 | 21.83 | 33.01 | 54.24 | ----- | ----- | ----- | |
| mean | $\bar{x}_1$ 33.65 | $\bar{x}_2$ 48.34 | $\bar{x}_3$ 55.38 | $\bar{x}_4$ 60.82 | $\bar{x}_5$ 84.8 | $\bar{x}_6$ 318.4 | $\beta=\sum_{k=1}^{6}\bar{x}_k=601.3$ |

وبما أنّه لدينا ستة فصول (وفقاً لهذا المثال) فإنّه سيكون لدينا $\alpha=6\times 100=600$ ، ومن ثمّ نجد أنّ $\gamma=\frac{\alpha}{\beta}=\frac{600}{601.3}=0.998$، وهكذا يكون لدينا مُقدّرات المركّبات الموسميّة هي:

$$\hat{S}_1=\gamma\times 33.65=0.998\times 33.65=33.583$$

$$\hat{S}_2=\gamma\times 48.34=0.998\times 48.34=48.243$$

$$\hat{S}_3=\gamma\times 55.38=0.998\times 55.38=55.269$$

$$\hat{S}_4=\gamma\times 60.82=0.998\times 60.82=60.698$$

$$\hat{S}_5=\gamma\times 84.8=0.998\times 84.8=84.630$$

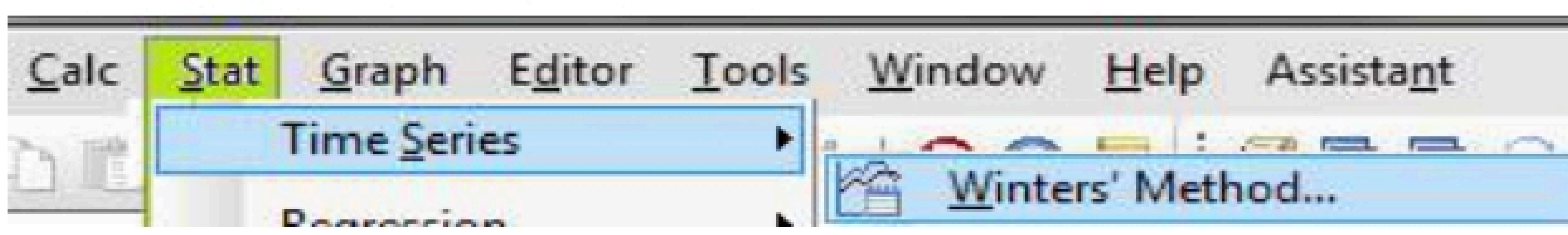
$$\hat{S}_6=\gamma\times 318.4=0.998\times 318.4=317.763$$

بهذا نأتي على نهاية هذه الفقرة، ونقدّم فيما يلي إحدى الطرائق المستخدمة في توليد التنبؤات من بيانات متسلسلة زمنية.

## (١٢,٧) طريقة وينتر لتوليد تنبؤات من متسلسلة زمنية

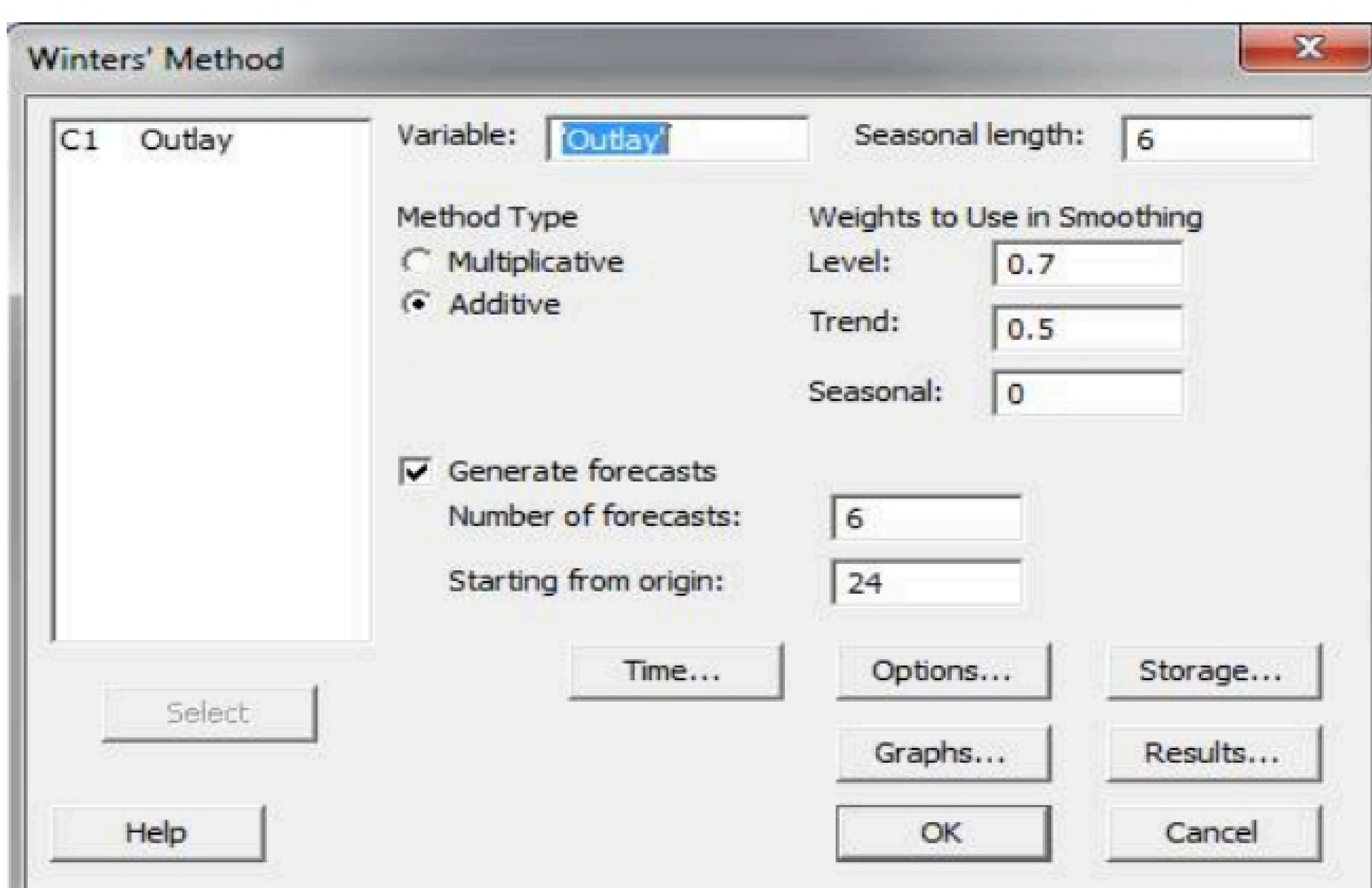
**Winter's Method for Generate Forecasts of a Time Series**

قبل أن نختم هذا الفصل سنقوم باستخدام طريقة وينتر لتوليد بعض التنبؤات للمتسلسلة الزمنية المُعطاة في المثال السابق، وذلك باستخدام الخطوات الآتية:




**Stat → Time Series → Winters' Method.**

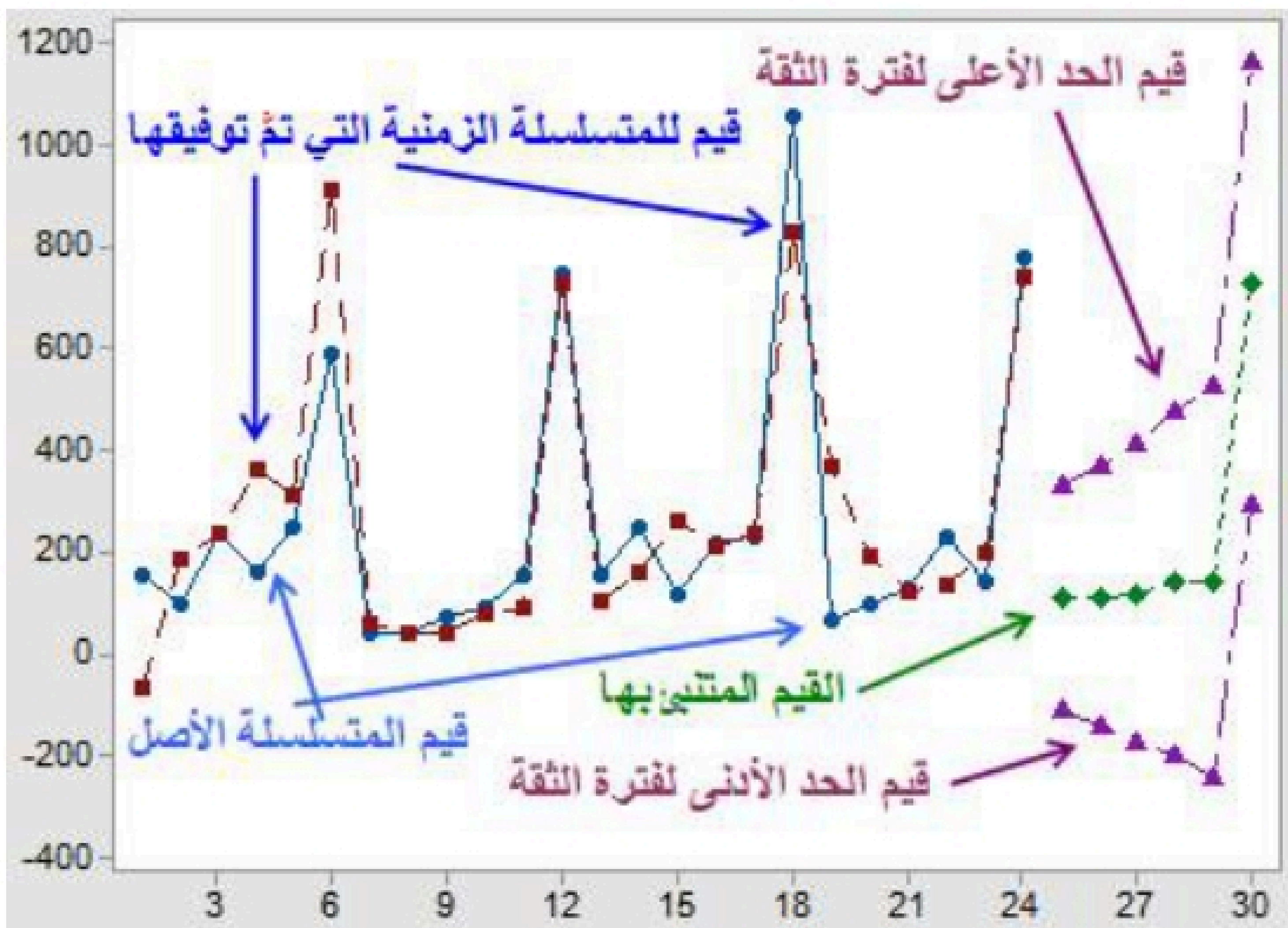
ومن ثمَّ في نافذة Variable أدخل كلمة Outlay وفي نافذة Seasonal Length أدخل العدد 6، ومن أجل النموذج في Model Type نفعِّل Additive. أمّا الأوزان α و γ و δ فيمكن التعديل عليها بحيث نحصل على أصغر خطأ معياري MSD ممكن للتقدير (وكما ذكرنا سابقاً بأنَّ هذه العملية تحتاج إلى خبرة وجهد للوصول إلى القيم المناسبة لهذه الأوزان)، ومن أجل مثالنا هذا اخترنا قيم الأوزان كما في التقرير الآتي الذي يقدِّم لنا 95% فترة ثقة للتنبؤات المولَّدة أيضاً، وذلك بعد تفعيل Generate forecasts وإدخال 6 في نافذة Number of forecasts، وكذلك 24 في نافذة Starting from origin، وأخيراً الضغط على أيقونة **OK**.




فنحصل على التقرير الآتي (سنضعه في جدول بعمودين مع أخذ ما المُعطيات التي نحتاجها فقط):

| Winters' Method for Outlay | Forecasts | | | |
|---|---|---|---|---|
| Additive Method | **Period** | **Forecast** | **Lower** | **Upper** |
| Data Outlay | **25** | **111.839** | **-110.318** | **334.00** |
| Length 24 | **26** | **112.830** | **-142.182** | **367.84** |
| Smoothing Constants | **27** | **116.821** | **-178.102** | **411.74** |
| α (level) 0.6 | **28** | **138.811** | **-200.597** | **478.22** |
| γ (trend) 0.3 | **29** | **142.052** | **-244.843** | **528.95** |
| δ (seasonal) 0.0 | **30** | **728.293** | **291.889** | **1164.70** |
| Accuracy Measures | | | | |
| **MSD** 17185.2 | | | | |

والعرض البياني المرافق لهذا التقرير يقدِّمه الشكل الآتي:




الشكل (٢٠, ١٢. هـ)

في الحقيقة إنَّ طريقة وينتر تُعدّ من الطرائق الجيدة في توفيق الكثير من المتسلسلات الزمنية، ولكن استخدامها يحتاج إلى خبرة وجهد، وفيما يلي نقدِّم موجزاً حول هذه الطريقة.

### (١, ٧, ١٢) طريقة وينتر في توفيق نموذج متسلسلة زمنية

إنَّ طريقة وينتر تقوم على أساس تنعيم البيانات بوساطة التنعيم الأسّيَ لـ Holt-Winters الذي يزودنا بقليل من التنبؤ متوسط المدى، حيث يمكننا استعمال هذا الإجراء عندما يكون لدينا اتجاه عام ومركّبة موسميّة (وكحالة خاصّة مركّبة دوريّة) في المتسلسلة الزمنية.

من أجل التحليل باستخدام طريقة وينتر يجب اختيار أحد نموذجين منطلقين من معرفتنا المسبقة لطول دور الموسم $p$ ، وهذين النموذجين (الممكن استخدام أحدهما في تنفيذ التحليل) هما:

١ - نموذج الجمع Additive Model، وقد تحدثنا سابقاً عن كيفية استخدام هذا النموذج.

أمّا العلاقات المستخدمة وفقاً لهذا النموذج من أجل الحصول على التقديرات للبيانات التي قيد الدراسة (حيث طول دور الموسم $p$ يجب أن يكون معلوماً )، فإنَّها تتمّ بحساب المقادير الآتية أولاً :

$$L_t = \alpha\left(Y_t - S_{t-p}\right) + (1-\alpha)\left(L_{t-1} + T_{t-1}\right)$$

$$T_t = \gamma\left(L_t - L_{t-1}\right) + (1-\gamma)\, T_{t-1}$$

$$S_t = \delta\left(Y_t - L_t\right) + (1-\delta) S_{t-p}$$

علماً أنَّ $L_t$ هي قيمة المستوى (أو التمهيد) Level عند اللحظة الزمنية $t$، و $\alpha$ هو ثابت ويمثِّل الوزن للمستوى (أو الوزن للتمهيد) Weight for the level، و $T_t$ هي قيمة الاتجاه العام Trend عند اللحظة الزمنية $t$، و$\gamma$ هو ثابت يمثِّل الوزن للاتجاه العام Weight for the trend، و $S_t$ هي قيمة المركّبة الموسميّة Seasonal عند اللحظة الزمنية $t$، وأخيراً $\delta$ هو ثابت يمثِّل الوزن للمركّبة الموسميّة Weight for the seasonal component.

عندئذ تحسب التقديرات $\hat{x}_t$ وفقاً للعلاقة الآتية:

$$\hat{x}_t = L_{t-1} + T_{t-1} + S_{t-p}$$

علماً أنَّ $\hat{x}_t$ هي قيمة التنبؤ المناسبة للوحدة الزمنية التالية للحظة الزمنية $t-1$ ، بمعنى أنَّه وفقاً لهذه الطريقة يتمَّ التنبؤ فيها باستخدام التكرار التقدّمي للحصول على التنبؤات ولكن واحدة فواحدة فقط إلى حين الوصول إلى نهاية الفترة الزمنية المطلوب التنبؤ عنها.

٢- **نموذج الضرب** Multiplicative Model، وقد تحدثنا سابقاً عن كيفية استخدام هذا النموذج أيضاً، وأمَّا العلاقات المستخدمة وفقاً لهذا النموذج من أجل الحصول على التقديرات للبيانات التي قيد الدراسة (تحت الفرض أنَّ طول دور الموسم $p$معلوماً)، فإنَّها تتمّ بحساب المقادير الآتية أولاً:

$$L_t = \alpha\left(\frac{Y_t}{S_{t-p}}\right) + (1-\alpha)\left(L_{t-1} + T_{t-1}\right) \qquad \& \qquad T_t = \gamma\left(L_t - L_{t-1}\right) + (1-\gamma)T_{t-1}$$

$$S_t = \delta\left(\frac{Y_t}{L_t}\right) + (1-\delta)S_{t-p}$$

علماً أنَّ الرموز الواردة في هذه العلاقات لها الدلالات السابقة نفسها. عندئذ تحسب التقديرات $\hat{x}_t$ وفقاً للعلاقة الآتية:

$$\hat{x}_t = \left(L_{t-1} + T_{t-1}\right) \cdot S_{t-p}$$

وهنا $\hat{x}_t$ هي قيمة التنبؤ المناسبة للوحدة الزمنية التالية للحظة الزمنية $t-1$ أيضاً، بمعنى أنَّه وفقاً لهذه الطريقة يتمَّ التنبؤ فيها باستخدام التكرار التقدّمي للحصول على التنبؤات، ولكن واحدة فواحدة فقط إلى حين الوصول إلى نهاية الفترة الزمنية المطلوب التنبؤ عنها.

عند اختيار أي من النموذجين السابقين فإنَّ طريقة وينتر تقدِّم لنا حسابات فعاَّلة للتقدير باستعمال مَعَالم التنعيم الثلاثة: مَعْلمة المستوى Level Parameter ، مَعْلمة الاتجاه العام Trend Parameter وأخيراً المَعْلمة الموسميّة Seasonal Parameter، علماً أنَّ هذه المَعَالم تأخذ قيمها في الفترة $[0,1]$ .

إنَّ برنامج Minitab ينطلق بتقدير متوازن لهذه المَعَالم عند القيمة **0.2** ، وعلى الباحث أن يقوم بالتعديل على هذه المَعَالم تجريبياً (أو حسب خبرته) للحصول على التقدير الدقيق المطلوب للمسالة التي قيد الدراسة، بمعنى أنَّه لا توجد قيم مُثلى للمَعَالم الثلاثة السابقة (مَعْلمة المستوى، مَعْلمة الاتجاه العام، المَعْلمة الموسمي) وفقاً لطريقة وينتر بحيث يحصل المرء على التقديرات الأكثر دقَّةً كما هو الحال لدى ما يُعرف باسم **طريقة التنعيم الأسّي المفرد والمزدوج** (التنعيم الأسّي المفرد والمزدوج من الطرائق التي يمكن إنجازها في برنامج **Minitab**أيضاً).

الآن وبعد الوصول إلى التعديل النهائي على قيم مَعَالم المستوى، الاتجاه العام والموسمي يمكننا استخدام هذا النموذج للحصول على تقديرات جيدة للتنبؤات بعد تزويده بطول الموسم (طول الدور للحركة الاهتزازية الممثّلة للمركّبة الدوريّة والموسميّة).

### (١٢,٧,١,١) توفيق النموذج Fitting of Model

إنَّ طريقة وينتر تستثمر قيم المَعَالم الثلاثة (مَعْلمة المستوى، مَعْلمة الاتجاه العام، المَعْلمة الموسمي) في الحسابات ضمن كل دور من الأدوار الموجودة في مسار المتسلسلة الزمنية، وعلاوة على ذلك تستخدم هذه المَعَالم الثلاثة لتحسين وتطوير قيم المركّبات الموافقة لها في كل دور من الأدوار. إنَّ القيم الأولية لمركّبات الاتجاه العام والمستوى يُستحْصل عليها من الانحدار الخطي مع مرور الزمن، وأمَّا القِيَم الأولية للمركّبة الموسميّة فإنَّه يُستحْصل عليها من استعمال متغيّر انحدار زائف (وهمي) مستخدماً البيانات المقدَّرة للاتجاه العام النَاتجة عن الانحدار الخطي.

أمّا بالنسبة إلى جودة توفيق النموذج فإنَّه توجد معايير للحكم على مدى هذه الجودة، وهذا يعدّ من ضرورات الحصول على تنبؤات مقبولة ومنسجمة مع سلوك المتسلسلة الزمنية، وذلك لأنَّ أغلب التنبؤات ومهما كانت طريقة التنبؤ تميل إلى أن تكون غير صحيحة مئة بالمئة، ولذلك لا بدّ من تقييم جودة التنبؤ بمقارنة القيم الحقيقية بالقيم المقدرة، وهذه المقارنة ستكشف لنا حجم الأخطاء في التنبؤ (ومن ثمَّ تقييم جودة التنبؤ).

إنَّ أغلب مقاييس جودة التوفيق تعتمد على قيم الانحرافات بين القيم الفعليّة للمتسلسلة الزمنية الأصل والقيم المُقدَّرة لها، ومن المعايير المستخدمة في قياس جودة التوفيق Measuring Fitting Accuracy في برنامج Minitab هي:

١- متوسط القيم المطلقة للانحرافات Mean Absolute Deviation، ويُعطى هذا المقياس بالعلاقة الآتية:

$$\mathbf{MAD} = \sum_{t=1}^{n} \frac{|y_t - \hat{y}_t|}{n}$$

٢- متوسط القيم المطلقة للنسب المئوية للأخطاء (أو متوسط القيم المطلقة للنسب المئوية للانحرافات) Mean Absolute Percentage Error، ويعطى هذا المقياس بالعلاقة الآتية:

$$\mathbf{MAPE} = \frac{100}{n} \sum_{t=1}^{n} \frac{|y_t - \hat{y}_t|}{y_t}$$

٣- متوسط مربَّع الانحرافات Mean Square Deviation (والذي اعتمدناه سابقاً)، ويعطى هذا المقياس بالعلاقة الآتية:

$$\mathbf{MSD} = \sum_{t=1}^{n} \frac{(y_t - \hat{y}_t)^2}{n}$$

علماً أنَّ $y_t$ هي قيمة المشاهدة التي رُصدت عند اللحظة الزمنية $t$، وأمّا $\hat{y}_t$ فهي القيمة المقدَّرة لـ $y_t$ عند اللحظة الزمنية $t$ وفقاً للنموذج المعتمد في التوفيق، وأخيراً $n$ هو عدد المشاهدات (أو حجم العيِّنة) من مسار المتسلسلة الزمنية التي أخضعت للدراسة.

### (١٢,٧,١,٢) التنبؤ Forecasting

تَستعمل طريقةُ وينتر كل من المستوى، والاتجاه العام والمركّبات الموسميّة لتوليد التنبؤات. إنّ التنبؤ لـ $k$ لحظة زمنية قادمة بدءًا من لحظة زمنية مفترضة $t$ يقوم البرنامج بحساب $L_t + k \cdot T_t$، ومن ثَّم يجمع (أو يضرب من أجل النموذج الضربي) هذه المركّبة مع (أو يضرب بـ) المركّبة الموسميّة وذلك من أجل الدور نفسه من الدورة السابقة Previous Cycle. علاوةً على ذلك تستخدم طريقة وينتر البيانات من لحظة البدء المفترضة لتوليد التنبؤات المطلوبة.

### (١٢,٧,١,٣) اختيار الأوزان وفقاً لطريقة وينتر

يمكنك إدخال الوزن أو مَعْلمة التنعيم للمستوى وللاتجاه العام والمركّبات الموسميّة. إنّ الأوزانَ الافتراضية كما ذكرنا سابقاً هي 0.2، ولكن يمكننا إدخال أي قيمة تقع بين 0 والـ 1 حسبما تقتضي الحاجة لتحسين دقة نتائج النموذج. إلاَّ أنَّه بغض النظر عن المركّبات فإنَّ الأوزان الكبيرة تُؤدّي إلى تغييرات أكثر سرعة في المركّبات، في حين أنَّ الأوزان الصغيرة تؤدي إلى تغييرات أقل سرعة. إنَّ المركّبات تؤثّر تباعاً في القيم المنعّمة (أو المقدَّرة) والمُتنبأ بها أيضاً، ولهذا فإنَّه يوصى عادة باستخدام أوزان صغيرة من أجل المتسلسلات الزمنية ذات مستوى ضوضاء عالٍ حول الإشارة أو الشكل، وأمّا الأوزان الكبيرة فإنَّه يوصى بها عادةً من أجل المتسلسلات الزمنية ذات مستوى ضوضاء منخفض حول الإشارة أو الشكل.

**هذا ما تيسَّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل**

# تمارين الفصل الثاني عشر

١- تحدث عن الأسباب التي تدفعنا لتحليل متسلسلة زمنية.

٢- تحت الفرض أنّنا قمنا بتحليل متسلسلة زمنية وقمنا بتعيين نموذج ملائم لها، فعندئذ علّل سبب ثقتنا بقبول تنبؤات هذا النموذج.

٣- ما هو سبب وجود المركّبة العشوائية في نموذج المتسلسلة الزمنية، وكيف تفسِّر الافتراض من أنَّ هذه المركّبة من طبيعة الضجيج الأبيض.

٤- ما هي الغاية من استخدام طريقة النسبة إلى متوسط متحرك في دراسة الأثر الموسمي.

٥- لتكن لدينا البيانات الآتية لعينة مأخوذة من مسار متسلسلة زمنية:

| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 63 | 101 | 84 | 116 | 110 | 137 | 100 | 94 | 129 | 98 |

والمطلوب ما يلي:

أ - تعيين الاتجاه العام الخطي لهذه المتسلسلة الزمنية.

ب- توليد متسلسلة زمنية أخرى من المتسلسلة السابقة باستخدام متوسطات متحركة بطول يساوي 3، ومن ثَّم رسم بصمة المتسلسلة الزمنية الناتجة عن المتوسطات المتحركة، قارن الشكل الناتج مع البصمة للمتسلسلة الزمنية الأصل، ماذا تلاحظ.

ج- تعيين معادلة كثيرة الحدود المناسبة للاتجاه العام لهذه المتسلسلة الزمنية باستخدام طريقة الفروق المتتالية.

٦- لتكن لدينا البيانات الآتية لعينة مأخوذة من مسار متسلسلة زمنية:

| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 12 | 22 | 18 | 14 | 27 | 16 | 13 | 26 | 18 | 29 |

والمطلوب ما يلي:

أ - رسم بصمة هذه المتسلسلة الزمنية.

ب- حساب معامل الخشونة لهذه المتسلسلة الزمنية.

ج- توليد متسلسلة زمنية من المتسلسلة السابقة باستخدام متوسطات متحركة بطول يساوي 2.

٧- لتكن لدينا البيانات الآتية لعينة مأخوذة من مسار متسلسلة زمنية:

| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |
|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 3 | 6 | 10 | 16 | 25 | 32 | 40 | 46 | 50 | 53 |

والمطلوب ما يلي:

أ - رسم بصمة هذه المتسلسلة الزمنية.

ب- استخدم الدالَّة المنطقية في تعيين الاتجاه العام لهذه المتسلسلة الزمنية.

ج- تنبأ بالقيمتين الموافقتين للفجوتين الزمنيتين 11 و12.

٨- عيِّن كثيرة الحدود المقدَّرة للاتجاه العام للمتسلسلة الزمنية الآتية:

| $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 1 | 5 | 8 | 14 | 26 | 35 | 50 | 64 | 82 | 98 | 120 | 145 |

٩- لتكن لدينا المشاهدات المقدَّمة في الجدول الآتي والمأخوذة من مسار متسلسلة زمنية تصف أرباح مضارب في البورصة (مقدَّرة ب 1000 وحدة نقدية) في كل شهر وعلى مدى 28 شهراً:

| الشهر $t$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| $x_t$ | 3 | 1 | 2 | 4 | 6 | 7 | 5 | 3 | 2 | 5 | 9 | 11 | 8 | 4 |
| الشهر $t$ | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 | 26 | 27 | 28 |
| $x_t$ | 3 | 6 | 11 | 15 | 12 | 8 | 4 | 8 | 13 | 17 | 15 | 7 | 5 | 8 |

والمطلوب ما يلي:

أ - ارسم بصمة هذه المتسلسلة الزمنية، ومن ثمَّ بيّن أيًّا من نموذجي الجمع أو الضرب أكثر توافقاً مع هذه المتسلسلة الزمنية مع التعليل.

ب- عيّن الاتجاه العام الخطي لهذه المتسلسلة الزمنية.

ج- بيّن هل كانت الأرباح الشهرية لهذا المضارب مستقلَّة أم لا؟

د- عيّن الدالَّة المقدِّرة للمركِّبة الدوريّة لهذه المتسلسلة الزمنية.

هـ- ولِّد بعض التنبؤات من النموذج الناتج.

١٠- لتكن لدينا البيانات المقدَّمة من خلال الجدول الآتي الذي يمثِّل عيِّنة مأخوذة من مسار متسلسلة زمنية لكمية منتجات زراعية مصدَّرة من إحدى المدن إلى المدن الأخرى في نهاية كل فصل من الفصول الأربعة من العام (مقدَّرةً بالطن):

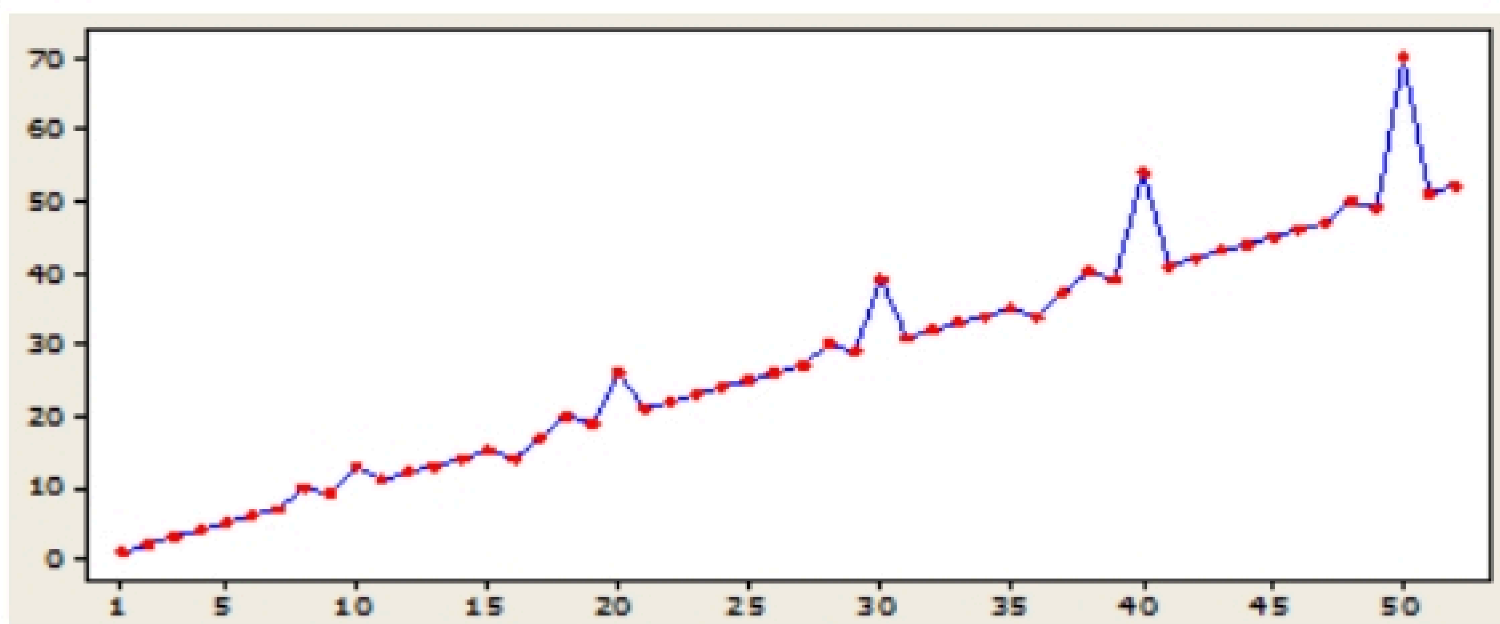
| | | الفصل | | | |
|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 |
| العام | 2010 | 125 | 225 | 143 | 133 |
| | 2011 | 100 | 320 | 150 | 135 |
| | 2012 | 111 | 285 | 135 | 150 |
| | 2013 | 120 | 240 | 155 | 140 |
| | 2014 | 115 | 275 | 135 | 125 |
| | 2015 | 105 | 310 | 120 | 138 |

والمطلوب ما يلي:

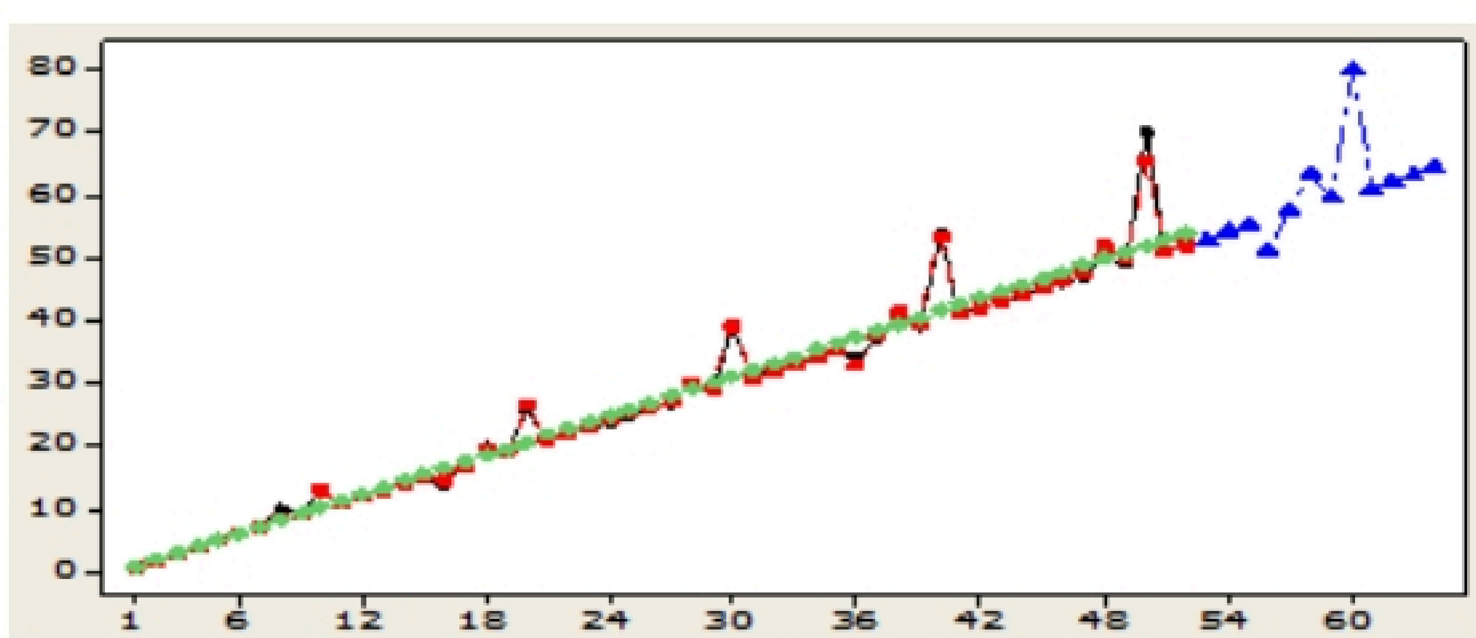
أ - رسم بصمة هذه المتسلسلة الزمنية.

ب- تقدير المركِّبات الموسميَّة لهذه المتسلسلة الزمنية باستخدام طريقة النسبة إلى متوسط متحرك.

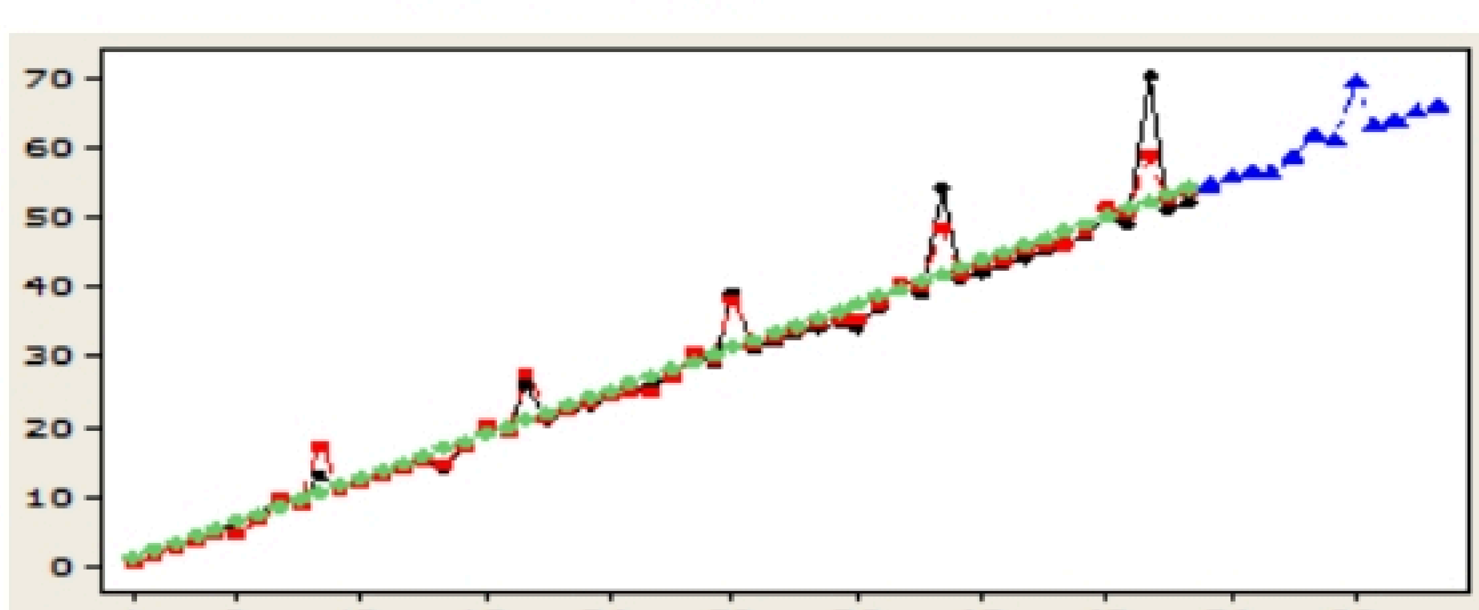
١١- لتكن لدينا العروض الآتية لمتسلسلة زمنية وقد تمَّ توفيق نماذج لها وفقاً لطريقة التجزئة (أمعن النظر في هذه الأشكال) مع الاتجاه العام والتنبؤات وفقاً لهذه النماذج:




المتسلسلة الزمنية الأصل




توفيق للمتسلسلة الأصل مع تنبؤات لدور كامل




توفيق للمتسلسلة الأصل مع تنبؤات لدور كامل

والمطلوب هو تحديد أيّ النموذجين الأخيرين هو نموذج جمع كتوفيق للمتسلسلة الأصل، وكيف تفسِّر استنتاجك؟

# الفصل الثالث عشر

## الأرقام القياسية
## INDEX NUMBERS

### (١٣,١) مفهوم الرقم القياسي
### Concept of Index Number

في الحقيقة يعود استخدام الأرقام القياسيّة إلى أكثر من قرنين من الزمن، حيث استخدمها الاقتصادي والمؤرخ الإيطالي كارلي Gian Rinaldo Carli (1720-1795) عام 1764 لمقارنة الأسعار في إيطاليا لسنة 1750 بالأسعار في سنة 1500، وكان استخدام الأرقم القياسيّة آنذاك حكراً على الاقتصاديين، ولكن بسبب التطبيقات الواسعة لهذا النوع من العلوم تمَّ كسر هذا الاحتكار بحيث أصبح وسيلة في يد كلّ مَن يريد استخدام الرقم القياسي، ومن ثمَّ شاع استخدامها، وتطور مفهومها بحيث أصبح هناك العديد من المقاييس في هذا المجال سنأتي على ذكر بعضها لاحقاً.

في الواقع أصبحت الأرقام القياسيّة مستخدمة في الكثير من القطّاعات العلميّة وعلى وجه الخصوص من قبل دارسي علم الاقتصاد، والاجتماع، والسياسة، والإدارة، والزراعة، و...، وذلك من أجل إجراء مقارنات وقياس التغيُّرات التي تطرأ في مجالات أعمالهم. هذا من جانب، ومن جانب آخر فقد تنوّعت استخدامات الأرقام القياسيّة، فنجد على سبيل المثال: الرقم القياسي لأسعار الجملة، والرقم القياسي للصادرات، والرقم القياسي للواردات، والأرقام القياسيّة للإنتاج الزراعي، والأرقام القياسيّة للإنتاج الصناعي، وتكاليف المعيشة. كذلك يمكننا بمعونة الأرقام القياسيّة مقارنة متوسط سعر عدَّة مواد في سنة واحدة مع متوسط سعر الكميّة نفسها من ذات المواد في عدد من السنوات المختلفة الأخرى. هكذا نجد أنَّ الرقم القياسي (للأسعار مثلاً) يُفيدنا في فهم وتفسير تغيّر الظروف الاقتصادية والتجارية مع مرور الزمن.

**ما هو الرقم القياسي؟**

إنَّ الرقم القياسي هو عبارة عن مؤشّر إحصائي يقيس التغيُّر النسبي الذي يَطرأ على ظاهرة معيَّنة (مثل: السعر، أو الكمية، أو القيمة، أو النشاط، أو الدرجات، أو ....) بالنسبة إلى أساس مُعيَّن قد يكون فترة زمنيّة معيَّنة أو مكانا جغرافيا محدَّداً أو مؤسّسة أو...، حيث تؤخذ قيمة الظاهرة كأساس لحساب الرقم القياسي، ويُطلق على الوقت أو المكان الذي تنسب إليه الظاهرة اسم **دورة الأساس** Base Period، وفي المراجع العربية تتنوع الترجمات لهذه العبارة، فإذا كانت الدراسة في مجال الدراسات الاقتصادية والتجارية، فإنَّه غالباً ما يُطلق عليها اسم **فترة الأساس** أو **عام الأساس** أو... لارتباطها بعامل الزمن، وفي حال كانت الدراسة في مجال الدراسات الطبية والنفسية، فإنَّه غالباً ما يُطلق عليها اسم **حالة الأساس** أو **الوضع الأساس** أو... لارتباطها بالحالة المرضية التي قيد الدراسة. أمّا إذا كانت الدراسة في مجال الدراسات الجغرافية أو المتعلّقة بمفاهيم جغرافية، فإنَّه يُطلق عليها اسم **المكان الأساس** أو **الموقع الأساس**. لكنَّنا سنعتمد في هذا الفصل تسميّة **دورة الأساس** مع ذكر بعض المسميّات المرافقة لها بين الحين والآخر. علماً أنَّ دورة الأساس هنا قد تعني:

– **شركة أو مؤسسة أو مدرسة** أو...

– **فترة زمنيّة محدَّدة**، ومن الممكن أن يكون طول هذه الفترة الزمنيّة ثوانيَ (كأن تكون الظاهرة التي قيد الدراسة هي عدّ جسيمات ألفا المنطلقة من مادة مشعَّة- كتجربة رذر فورد جايجر مثلاً)، ومن الممكن أن تمتد إلى ساعات أو أعوام أو قرون من الزمن (كما في دراسة قضايا الحروب).

– **حيِّز ما من الفضاء**، وهذا الحيِّز قد يكون خطيًّا كأن يكون سلك الكهرباء للتوتر العالي المعلَّق بين عامودين (كأن تكون الظاهرة التي قيد الدراسة هي قياس المسافة التي يصل إليها المجال الكهراطيسي المولَّد من مرور تيار كهربائي بجهد 100 كيلو فولت في هذا السلك في منطقتين مختلفتين)، ومن الممكن أن يكون لهذا الحيِّز بُعد ثنائي كأن تكون المساحة التي يُبنى عليها منزل أو بلدة أو...، ومن الممكن أن يكون ثلاثي البُعد كأن تكون طبقة الجو المحيطة بالأرض التي يُسمح فيها بالمرور للطيران المدني. هكذا نلاحظ أنَّ هذا الحيِّز يمكن أن يكون صغيراً جداً كما في الدراسات الميكروسكوبية (قد يصل إلى الميكرو والنانو متر) أو قد يكون كبيراً جداً كما في الدراسات المايكروسكوبية (قد يصل إلى آلاف وملايين الكيلومترات).

أمّا الشيء الذي ستنسب بياناته إلى بيانات دورة الأساس (التي يجب أن تكون من ذات الطبيعة) فإنَّها تُدعى **دورة المقارنة** Compare Period أو **الدورة الحالية (أو الدورة الجارية)** Current Period، وبما أنَّ البيانات التي ستنسب بعضها إلى بعض من ذات الطبيعة فإنَّ الحوار السّابق بخصوص دورة الأساس يبقى ساري المفعول على دورة المقارنة، والجمل الآتية توضِّح لنا العلاقة النسبية لدورة المقارنة مع دورة الأساس:

١ – نرغب في معرفة مدى التغيّر الذي طرأ على أسعار الفاكهة بين عامي 2004 و2014، فيكون عام 2004 **دورة الأساس** في حين يصبح عام 2014 هو **دورة المقارنة**.

٢– نرغب في معرفة مدى التغيّر الذي طرأ على النشاط الإشعاعي لمادّة مشعّة (من خلال معرفة شدّة وكمية الجسيمات المنطلقة من هذه المادة) بين عشر ثوان سبقت والآن، فتكون الثانية التي مرّت قبل عشرة ثوان من الآن هي **دورة الأساس** في حين تصبح الثانية التي نقيس خلالها عدد الجسّيمات المنطلقة من المادة المشعّة هي **دورة المقارنة**.

٤– نرغب في معرفة مدى التغيّر الذي طرأ على عدد ضحايا الحروب في القرن الرابع عشر وما يقابلها من ضحايا في القرن العشرين (القرن الماضي)، فيكون القرن الرابع عشر هو **دورة الأساس** في حين يصبح القرن العشرين هو **دورة المقارنة**.

٥– نرغب في معرفة مدى التغيّر الحاصل لكميات بعض المنتجات لشركة منتجة لقطع سيارات هونداي في اليابان مع شركة أخرى للمنتجات نفسها في الهند، فعندئذ الشركة المنتجة لقطع سيارات هونداي في اليابان تُؤخذ **كدورة أساس** في حين أنَّ الشركة المنتجة لقطع سيارات هونداي في الهند تلعب دور **دورة المقارنة**.

٥– نرغب في معرفة مدى التغيّر الحاصل لنشاط جراثيم (من خلال معرفة سرعة التكاثر ونفث السموم الناتجة عن الجراثيم-...) في مزرعة جرثومية بحجم 15 سنتيمترًا مكعبًا، والجراثيم الأخرى من ذات النوع التي تعيش في الهواء الطلق، فعندئذ المزرعة ذات حجم 15 سنتيمترًا مكعبًا تُؤخذ **كدورة أساس** في حين أنَّ الفضاء الذي تعيش فيه الكائنات الحيّة على الأرض هو **دورة المقارنة**.

هكذا نجد أنَّ الرقم القياسي هو مؤشِّر إحصائي يوضِّح المقارنة النسبية بين رقمين أحدهما يُستخدم كأساس للمقارنة. بعبارة أخرى، فإنَّ الرقم القياسي يقيس التغيّر النسبي لقيمتين لأشياء قيد الدراسة (صيغة الجمع هنا – لأشياء– إنَّما للتعميم ولا تفيد الحصر، بمعنى أنَّه قد تكون الدراسة معنيّة بشيء واحد فقط أو شيئين أو أكثر من ذلك) من:

– فترة زمنيّة إلى أخرى، علماً أنَّ إحدى هاتين القيمتين تمثِّل الفترة الزمنيّة المعنية (دورة الأساس) لهذا الشيء، والقيمة الأخرى تمثِّل هذا الشيء في الفترة الزمنيّة الأخرى (دورة المقارنة).

- أو من مكان إلى آخر، وفي هذه الحالة تكون إحدى هاتين القيمتين تمثِّل المكان المعني (دورة الأساس) لهذا الشيء، والقيمة الأخرى تمثِّل هذا الشيء في المكان الآخر (دورة المقارنة).

- أو من حالة إلى أخرى، وهنا تكون إحدى هاتين القيمتين تمثِّل الحالة المعنّية (دورة الأساس) لهذا الشيء، والقيمة الأخرى تمثِّل هذا الشيء لدى الحالة الأخرى (دورة المقارنة).

**أخيراً نشير إلى الآتي:**

١- من الممكن أن يكون لدينا مُعطيات عديدة تمثِّل كل واحدة منها فترة زمنيّة (أو مكان أو حالة)، فعندئذ ليس بالضرورة أن تكون الفترة الزمنيّة الأقل (أو المكان الأقرب أو الحالة الأدنى) هي فترة الأساس، وكذلك ليس بالضرورة أن تؤخذ كفترة مقارنة، بل من الممكن اختيار أيّ من هاتين الفترتين الزمنيّتين (أو المكانين أو الحالتين) كفترة أساس أو مقارنة وذلك حسب طبيعة المسألة وشروطها ومتطلباتها، فلو افترضنا أنَّه لدينا المعطيات الآتية لمتوسط أسعار كيلو اللحم في بلد ما مقدَّراً بوحدات نقديّة:

الجدول (١, ١٣)

| متوسط سعر كيلو اللحم<br>Price -year | العام<br>year |
|---|---|
| 10 | 1995 |
| 13 | 2000 |
| 18 | 2005 |
| 25 | 2010 |
| 40 | 2015 |

فيمكننا هنا أن نأخذ أيًا من هذه الأعوام على أنَّه دورة الأساس، فإذا كان اهتمامنا منصباً على مقارنة الأسعار إلى ما كانت عليه في عام 2000 مثلاً، فعندئذ يكون عام 2000 هو دورة الأساس، وأيّ عام آخر ستنسب قيمته إلى القيمة 13 (لعام 2000) سيلعب دور دورة المقارنة.

٢- نشير إلى أنَّه عند اختيار دورة الأساس ينبغي النظر بعناية فائقة إلى الأمور الآتية:

أ- ألا تكون دورة الأساس من الماضي البعيد جداً عن دورة المقارنة.

ب- أنَ تكون دورة الأساس والمقارنة من ذات الطبيعة، بمعنى أنّه لا يجوز هنا أن نستخدم شهرًا لدورة الأساس ويومًا لدورة المقارنة.

ج- أنَ تكون وحدات القياس المستخدمة في دورتي الأساس والمقارنة من ذات الطبيعة أيضاً.

د- في حال الأخذ بالحسبان الفترات السابقة فإنَّه يجب التفكير مليًّا في تحديد أيّ الدورتين ستكون دورة أساس وبشكل خاص عند استخدام المؤشِّرات الموزونة (سنأتي على ذكرها لاحقاً).

لقد لاحظنا أنَّ الأرقام القياسيّة تشغل مكاناً بارزاً بين المقاييس الإحصائية المستخدمة لدراسة الظواهر الاجتماعية، والاقتصادية، والعلمية، لا بل يؤكد الاقتصاديون أنَّ الأرقام القياسيّة تعدَّ وسيلة فعَّالة ومهمّة جداً من وسائل الإحصاء الحديث المستخدمة في الاقتصاد.

بهذا الشكل يكون الطريق قد مُهِّد جيداً لتعريف الرقم القياسي الذي سيقدِّمه لنا التعريف الآتي.

### (١ , ١ , ١٣) تعريف (الرقم القياسي Index Number)

إنَّ الرقم القياسي هو عدد في صيغة نسبية يُستخدم لقياس التغيُّر النسبيّ لظاهرة ما (مثل: السعر، الكمية، الحرارة، و....) أو لعدّة ظواهر من فترة زمنيّة إلى أخرى، أو من مكان إلى آخر، أو من حالة إلى أخرى.

### (١ , ١ , ١ , ١٣) ملاحظات

١ – إنَّ الرقم القياسي يُدعى مؤشّراً أيضاً، وهذا ما سنستخدمه لاحقاً في هذا الفصل.

٢ – في الواقع للأرقام القياسيّة أوجه مختلفة في المجتمعات الإحصائية، فقد تصادفنا تعاريف أخرى قد تختلف عن ذلك الذي قدِّم في التعريف السابق، فعلى سبيل المثال قد يعرِّف البعض الرقم القياسي على أنّه معدَّل المتسلسلة الزمنيّة كما يرد في بعض الدراسات للتغيّرات الموسميّة في المتسلسلات الزمنيّة، في حين يرى البعض الآخر أنَّ الرقم القياسي هو قيمة نسبية تُظهر بشكل مباشر التغيُّر المتوسط في الظواهر الاجتماعية كما يرد في بعض دراسات علم الاجتماع. هكذا يُلاحظ أنَّ دور الرقم القياسي في هذه الحالات إنَّما يقتصر على وصف إجمالي التغيُّر في الظاهرة ولا يتعداها إلى حالة التعميم (وهذه التعاريف للرقم القياسي تنضوي تحت الدراسات التقليدية)، ولكن الدراسات الحديثة تولي الرقم القياسي بعداً أكثر شمولية ممّا سبق، فهي ترى أنَّ الرقم القياسي يجب ألّا يقتصر عمله على الوصف الإجمالي للتغيُّر في الظاهرة فقط، وإنّما يجب أن يصف دور كلٍّ من العوامل التي أدت إلى إحداث هذا التغيُّر الإجمالي أيضاً سواء في دورة الأساس أو دورة المقارنة.

بناءً على ما سبق سيكون لدينا نماذج مختلفة من الأرقام القياسيّة، وسنبدأها بالمؤشّر الآتي.

## (١٣,٢) الأرقام القياسية البسيطة

**Simple Index Numbers**

في هذا الفصل سنقدِّم العديد من الأرقام القياسيّة (المؤشّرات) مبتدئين بالبسيط منها ومن ثَمَّ ننتقل إلى الأعم فالأكثر عموميّة.

إنَّ الأرقام القياسيّة التي تنشأ عن شيء واحد فقط تُسمَّى أرقاماً قياسيّةً بسيطة، وهذا يعني أنَّ الرقم القياسي البسيط يقيس التغيُّر النسبي بمتغيِّر واحد فقط، والتعريف الآتي يوضّح لنا هذا المفهوم.

### (١ , ٢ , ١٣) تعريف (الرقم القياسي البسيط Simple Index Number)

لتكن $p_o$ قيمة تمثِّل شيئًا (واحدًا فقط) في دورة الأساس، وكذلك $p_c$ قيمة تمثِّل ذلك الشيء نفسه في دورة المقارنة، فعندئذ تُدعى النسبة المئوية الآتية (التي سنرمز لها بـ $\mathbf{I}_S$) رقماً قياسياً بسيطاً:

$$\mathbf{I}_S = \frac{p_c}{p_o} \times 100\ \% \qquad [13,1]$$

### (١ , ١ , ٢ , ١٣) ملاحظات

١ – من الواضّح أنَّ الرقم القياسي البسيط (المؤشّر البسيط) هو عدد مجرَّد من وحدة القياس، وذلك لأنّها ناتج قسمة مقدار على مقدار آخر له وحدة القياس نفسها؛ لأنَّ البيانات المقسومة على بعضها من ذات الطبيعة (ريال على ريال، أو كغ على كغ، أو ...).

٢ – إنَّ قيمة الرقم القياسي البسيط لأي شيء من دورة الأساس يساوي 100%، ومن ثَمَّ قيمة الرقم القياسي البسيط المحسوب بالعلاقة السابقة [13,1] خاص بدورة المقارنة.

٣ – إنَّ الرقم القياسي البسيط $\mathbf{I}_S$ هو قيمة نسبية Relative Value تقيس التغيّر النسبي لقيمتين لشيء واحد قيد الدراسة من

فترة زمنيّة إلى أخرى، أو من مكان إلى آخر أو من حالة إلى أخرى، وكذلك يُلاحظ أنَّ الأشياء في هذا النوع من الدراسة تُدرس كلًّا على حدة. لكن إذا كان لدينا أكثر من شيء واحد خاضع للدراسة نفسها فإنَّه يُنصح بجمع النتائج في جدول واحد.

٤- عندما يكون الرقم القياسي البسيط $\mathbf{I}_S$ ناتج عن دراسة الأسعار، فعندئذ يُدعى $\mathbf{I}_S$ **مؤشِّر السعر البسيط** The Simple Price Index.

٥- إذا كانت قيمة المؤشّر $\mathbf{I}_S$ أكبر من 100، فإنّنا نطرح 100 من قيمة هذا المؤشّر، والنتيجة هي نسبة الزيادة في قيمة الشيء الذي هو قيد الدراسة من دورة الأساس إلى دورة المقارنة، وتقرأ كنسبة مئوية. وأمّا إذا كانت قيمة المؤشّر $\mathbf{I}_S$ أقل من 100، فعندئذ نطرح قيمة هذا المؤشّر من 100 والنتيجة هي النسبة المئوية لذلك النقص في تكلفة الشيء الذي قيد الدراسة في دورة الأساس ممّا هي عليه في دورة المقارنة، وتقرأ كنسبة مئوية أيضاً.

٦- بفرض أنَّه لدينا $k$ شيء $T_1$، $T_2$،...و $T_k$ وكان $\mathbf{I}_{S,1}$، $\mathbf{I}_{S,2}$،...و $\mathbf{I}_{S,k}$ هي الأرقام القياسيّة البسيطة (المؤشّرات البسيطة) لتلك الأشياء على الترتيب، فعندئذ يُقال عن المقدار الآتي إنَّه **المعدَّل البسيط** Simple Average للقيم النسبية لتلك الأشياء:

$$\mathbf{SA} := \frac{1}{k}\sum_{i=1}^{k}\mathbf{I}_{S,i} \qquad [13,2]$$

وعندما تكون الدراسة متعلّقة بالأسعار فإنَّه يُدعى **المعدَّل البسيط للقيم النسبية للأسعار** Simple Average of Price Relatives، ويقرأ الناتج كنسبة مئوية.

٧- في دراستنا المقبلة:

أ- سنرمز بـ $p_{o,1}$، $p_{o,2}$،...و $p_{o,k}$ للقيم الناتجة عن دراسة الأشياء $T_1$، $T_2$،...و $T_k$ في دورة الأساس على الترتيب (استخدامنا للحرف ($o$) في دليل القيم جاء من كلمة old، للدلالة على دورة الأساس بمعنى الدورة القديمة).

ب- سنرمز بـ $w_{o,1}$، $w_{o,2}$،...و $w_{o,k}$ لأوزان (أو تكرارات) القيم $p_{o,1}$، $p_{o,2}$،...و $p_{o,k}$ على الترتيب.

ج- سنرمز بـ $p_{c,1}$، $p_{c,2}$،...و $p_{c,k}$ للقيم الناتجة عن دراسة الأشياء $T_1$، $T_2$،...و $T_k$ في دورة المقارنة على الترتيب (استخدامنا للحرف (c) في دليل القيم جاء من كلمة Compare، للدلالة على دورة المقارنة أو بمعنى الدورة الجارية (أو الحالية) **Current Period**).

د- سنرمز بـ $w_{c,1}$، $w_{c,2}$،...و $w_{c,k}$ لأوزان (أو تكرارات) القيم $p_{c,1}$، $p_{c,2}$،...و $p_{c,k}$ على الترتيب.

### (١٣,٢,١,٢) أمثلة

١- لنفترض أنَّ أسرةً تهتمّ بموضوع تكلفة استهلاكها السنوي للمواد الغذائية (مقدَّرة ب **100 كيلوغرام**) المتمثِّلة بالفاكهة، والخضراوات، والدقيق ومشتقاته، واللحوم، وأنَّ متوسط سعر الكيلوغرام الواحد لكلّ نوع من هذه السّلع الغذائية (مقدَّراً بـ وحدة نقدية **Monetary Units**) قد بدأ بقيمة ما في العام (1) Price-year (دورة الأساس)، علماً أنَّ العام (1) يمكن أن يشير إلى أي عام سابق. كذلك سنفترض أنَّ متوسط سعر الكيلو غرام الواحد (مقدَّرا بالوحدة النقدية نفسها) لكل نوع من هذه السّلع قد أصبح له سعر في عام تال آخر سنرمز له ب العام (2) (دورة المقارنة)، علماً أنَّ العام (2) ليس بالضرورة أن يكون العام التالي مباشرة للعام (1) وإنّما هو عام آخر يأتي بعد العام (1)، وكذلك ليس بالضرورة أن يكون العام (2) هو العام الحالي (أو العام الجاري) وإنَّما يمكن له أن يكون عامًا سابقًا للعام الحالي ولكنَّه بعد العام (1) حتماً. الآن لو افترضنا أنَّ أسعار هذه السّلع في العامين (1) و(2) مقدَّمة كما في الجدول الآتي:

الجدول (٢, ١٣ . أ)

| ر.م. $i$ | المنتج | دورة الأساس السعر في العام (1) Price -year | دورة المقارنة السعر في العام (2) Price -year |
|---|---|---|---|
| 1 | الفاكهة | $5 = p_{o,1}$ | $6 = p_{c,1}$ |
| 2 | الخضروات | $6 = p_{o,2}$ | $4 = p_{c,2}$ |
| 3 | الدقيق ومشتقاته | $2 = p_{o,3}$ | $2 = p_{c,3}$ |
| 4 | اللحوم | $20 = p_{o,4}$ | $33 = p_{c,4}$ |

فعندئذ لنقم بحساب قيم المؤشِّرات البسيطة (الأرقام القياسيّة البسيطة) لأسعار هذه السّلع الأربع، فنجد (بحسب العلاقة [13,1]) أنَّ المؤشِّر البسيط لسعر سلعة:

أ- الفاكهة يساوي:

$$\mathbf{I}_{S,1} = \frac{p_{c,1}}{p_{o,1}} \times 100 = \frac{6}{5} \times 100 = 120\%$$

ومن ثمَّ نسبة الزيادة في أسعار الفاكهة من دورة الأساس إلى دورة المقارنة يساوي 20%.

ب- الخضراوات تساوي $\mathbf{I}_{S,2} = \frac{4}{6} \times 100 = 67\%$ ، ومن ثمَّ نسبة النقصان في أسعار الخضروات من دورة الأساس إلى دورة المقارنة يساوي 33%.

ج- الدقيق ومشتقاته تساوي $\mathbf{I}_{S,3} = 100\%$، ومن ثمَّ لا وجود لأي تغيُّر في أسعار الدقيق ومشتقاته من دورة الأساس إلى دورة المقارنة.

د- اللحوم تساوي $\mathbf{I}_{S,4} = 165\%$، ومن ثمَّ نسبة الزيادة في أسعار اللحوم من دورة الأساس إلى دورة المقارنة تساوي 65%.

هكذا يمكننا تدوين نتائج هذه الدراسة في الجدول الآتي:

الجدول (٢, ١٣ . ب)

| ر.م. $i$ | المنتج | دورة الأساس السعر في العام (1) | مؤشِّر السعر في العام (1) Index-year 1 | دورة المقارنة السعر في العام (2) | مؤشِّر السعر في العام (2) Index-year 2 |
|---|---|---|---|---|---|
| 1 | الفاكهة | 5 | 100 % | 6 | 120 % |
| 2 | الخضراوات | 6 | 100 % | 4 | 67 % |
| 3 | الدقيق ومشتقاته | 2 | 100 % | 2 | 100 % |
| 4 | اللحوم | 20 | 100 % | 33 | 165 % |

ومنه نجد (بحسب العلاقة [13,2]) أنَّ قيمة مُعدَّل المؤشِّر البسيط لأسعار السّلع التي قيد الدراسة تساوي:

$$\mathbf{SA} := \frac{\sum_{i=1}^{4} \mathbf{I}_{S,i}}{4} = \frac{120+67+100+165}{4} = \frac{452}{4} = 113$$

ومن ثمَّ المعدَّل البسيط للسعر 113%، وهذا يعني أنَّ المعدَّل البسيط للقيم النسبية للسعر يساوي 113%، وبالتالي يوجد ارتفاع عام في أسعار هذه السّلع بمعدَّل 13% من العام (1) إلى العام (2).

٢- نعلم أنَّه في كل مرحلة ثانوية ثلاث مستويات دراسيّة (أول، وثان، وثالث ثانوي). الآن سنفترض أنَّنا نرغب في معرفة التغيّر الطارئ على تحصيل طلبة هذه المراحل بين مدارس المدن والأرياف، فقمنا باختيار عشوائي لإحدى المدن، ومن ثمَّ اختيار عشوائي لريف من أريافها، وبعد ذلك أحصيت النسب المئوية لأعداد الطلاب الذين حصلوا على معدَّل عام أكثر من 92% في نهاية العام الدراسي في كل مستوى من المستويات الثلاث، فكانت لدينا النتائج الآتية:

الجدول (١٣,٣ . أ)

| ر.م. $i$ | طلاب | دورة الأساس في المدينة المكان (1) | دورة المقارنة في الريف المكان (2) |
|---|---|---|---|
| **1** | **الأول الثانوي** | **65%** | **75%** |
| **2** | **الثاني الثانوي** | **52%** | **60%** |
| **3** | **الثالث الثانوي** | **83%** | **78%** |

لاحظ أنَّنا جعلنا في هذا الجدول المدينة هي دورة الأساس (مكان الأساس، ومن الممكن تسميتها المكان (1) أيضاً)، والريف التابع لهذه المدينة هو دورة المقارنة (مكان المقارنة، ومن الممكن تسميتها المكان (2) أيضاً). بالطبع كان من الممكن أن نأخذ العكس أيضاً حيث تُفسَّر النتائج بشكل معكوس في هذه الحالة. الآن لحساب المؤشّر البسيط لنسب الطلاب الذين حصلوا على مُعدَّل عام أعلى من 92% نجد (بحسب العلاقة [13,1]) أنَّ طلاب مستوى:

أ- الأول ثانوي يساوي:

$$\mathbf{I}_{S,1}=\frac{p_{c,1}}{p_{o,1}}\times 100=\frac{0.75}{0.65}\times 100=115.38\%$$

وهذا يعني أنَّ نسبة طلاب الأول ثانوي الذين على حصلوا على معدَّل عام في نهاية العام الدراسي أكثر من 92% من أبناء الريف تزيد على نسبة طلاب الأول ثانوي الذين على حصلوا على المعدَّل نفسه من أبناء المدينة بنسبة تساوي 15.38%.

ب- الثاني ثانوي يساوي $\mathbf{I}_{S,2}=\frac{0.60}{0.52}\times 100=115.38\%$، وهذا يعني أنَّ نسبة طلاب الثاني ثانوي الذين على حصلوا على معدَّل عام في نهاية العام الدراسي أكثر من 92% من أبناء الريف تزيد على نسبة طلاب الثاني ثانوي الذين حصلوا على المعدَّل نفسه من أبناء المدينة بنسبة تساوي 15.38% أيضاً.

ج- الثالث ثانوي يساوي $\mathbf{I}_{S,3}=93.98\%$، وهذا يعني أنَّ نسبة طلاب الثالث ثانوي الذين على حصلوا على معدَّل عام في نهاية العام الدراسي أكثر من 92% من أبناء الريف تقل عن نسبة طلاب الثالث ثانوي الذين على حصلوا على المعدَّل نفسه من أبناء المدينة بنسبة تساوي 6.02%.

وبوضع قيمة المؤشّر البسيط 100% لنسب الطلاب في المدينة، والذين على حصلوا على معدَّل عام في نهاية العام الدراسي أكثر من 92% في كل مستوى من المستويات الثلاث، فعندئذ يمكننا تدوين نتائج هذه الدراسة في الجدول الآتي:

الجدول (٣، ١٣. ب)

| ر.م. $i$ | طلاب | في المدينة المكان (1) | مؤشِّر معدَّل الطلاب في المدينة المكان (1) | في الريف المكان (2) | مؤشِّر معدَّل الطلاب في الريف المكان (2) |
|---|---|---|---|---|---|
| 1 | الأول الثانوي | 0.65 | 100 % | 0.75 | 115.38 % |
| 2 | الثاني الثانوي | 0.52 | 100 % | 0.60 | 115.38 % |
| 3 | الثالث الثانوي | 0.83 | 100 % | 0.78 | 93.98 % |

ومنه نجد (بحسب العلاقة [13,2]) أنَّ قيمة المعدَّل البسيط للقيم النسبية للنسب التي هي قيد الدراسة تساوي:

$$\mathbf{SA} := \frac{115.38+115.38+93.98}{3} = \frac{324.74}{3} = 108.25\%$$

ومن ثَمَّ المعدَّل البسيط للقيم النسبية للنسب التي قيد الدراسة يساوي 108.25%، وهذا يعني أنَّ الذين حصلوا على معدَّل عام في نهاية العام الدراسي أكثر من 92% من أبناء الريف يزيد على الذين حصلوا على المعدَّل نفسه من العام نفسه من أبناء المدينة بمقدار 8.25%.

٣- تقوم شركة تجارية بحملة إعلامية لترويج منتجاتها، ولذلك تدفع لوسائل الإعلام المختلفة مبلغاً من المال في بداية العام الذي تقرّر فيه بدء حملتها الإعلاميّة (القيم الآتية مقدَّرةٌ ب 1000 وحدة نقدية)، فإذا كانت المبالغ المدفوعة في عام 2010 و2014 لمؤسسة واحدة فقط من مؤسسات: الصحافة، والرائي، وإعلانات الطرق، والمذياع مقدَّمة كما في الجدول الآتي:

الجدول (٤، ١٣. أ)

| ر.م. $i$ | المؤسسة الإعلامية Media Foundation | مدفوعات المؤسسة في 2010 Price – year 2010 | مدفوعات المؤسسة في 2014 Price – year 2014 |
|---|---|---|---|
| 1 | Press الصحافة | 25 | 32 |
| 2 | Television الرائي | 65 | 78 |
| 3 | Roads ads إعلانات الطرق | 10 | 15 |
| 4 | Radio المذياع | 15 | 12 |

فعندئذ (بحسب العلاقة [13,1]) نجد أنَّ قيمة المؤشِّر البسيط في تكلفة:

أ- النشر في الصحافة يساوي:

$$\mathbf{I}_{S,1} = \frac{p_{c,1}}{p_{o,1}} \times 100 = \frac{32}{25} \times 100 = 128\%$$

ومن ثَمَّ نسبة الزيادة في تكلفة النشر في الصحافة من دورة الأساس (عام 2010) إلى دورة المقارنة (عام 2014) تساوي 28%.

ب- الإعلان في الرائي (التلفزيون) يساوي $\mathbf{I}_{S,2} = \frac{78}{65} \times 100 = 120\%$، ومن ثَمَّ نسبة الزيادة في تكلفة الإعلان في الرائي من دورة الأساس إلى دورة المقارنة تساوي 20%.

ج- الإعلانات الطرقية تساوي $\mathbf{I}_{S,3} = 150\%$، ومن ثَمَّ نسبة الزيادة في تكلفة الإعلانات الطرقية من دورة الأساس إلى دورة المقارنة تساوي 50%.

د- الإعلانات في المذياع (الراديو) تساوي $\mathbf{I}_{S,4} = 80\%$ ، ومن ثمَّ نسبة النقصان في تكلفة الإعلانات الإذاعية من دورة الأساس إلى دورة المقارنة تساوي 20%.

الآن بوضع قيمة المؤشِّر البسيط 100% لمدفوعات المؤسسة في العام 2010، فإنَّه يمكننا تدوين نتائج هذه الدراسة في الجدول الآتي (٣ , ١٣ . ب)، حيث نجد منه (بحسب العلاقة [13,2]) أنَّ قيمة المعدَّل البسيط للقيم النسبية لتكاليف الإعلام تساوي:

$$\mathbf{SA} := \frac{128+120+150+80}{4} = \frac{478}{4} = 119.5\%$$

ومن ثمَّ المعدَّل البسيط للسعر 119.5%، ومنه نلاحظ وجود ارتفاع عام في تكاليف الإعلانات في المؤسسات الأربع بمعدَّل 19.5% من العام 2010 إلى العام 2014.

الجدول (٤ , ١٣ . ب)

| ر.م. $i$ | المؤسسة الإعلامية | مدفوعات المؤسسة في العام 2010 | مؤشِّر التكلفة للعام 2010 | مدفوعات المؤسسة في العام 2014 | مؤشِّر التكلفة للعام 2014 |
|---|---|---|---|---|---|
| 1 | الصحافة | 25 | 100% | 32 | 128% |
| 2 | الرائي | 65 | 100% | 78 | 120% |
| 3 | إعلانات الطرق | 10 | 100% | 15 | 150% |
| 4 | المذياع | 15 | 100% | 12 | 80% |

إنَّ تفسير الزيادة أو النقص في مثل هذه المسائل ليس بالأمر البسيط غالباً، فعلى سبيل المثال إذا أخذنا موضوع الزيادة، فإنَّ ذلك قد يكون متعلِّقاً بجملة أسباب متداخلة منها على سبيل المثال لا الحصر: الغلاء الذي طرأ على أسعار المواد التي تستخدمها وسائل الإعلام خلال الفترة الزمنيّة 2010 إلى 2014، والغلاء العام للمعيشة الذي حصل ما بين عامي 2010 و 2014 والذي أجبر هذه المؤسسات على زيادة أجور عامليها ومن ثَمّ الزيادة في التكلفة العامّة للعمل، وقد يكون هناك أسباب أخرى كالضرائب الحكومية، و... ، بمعنى آخر إنَّ إظهار المؤشِّر للزيادة لا يعني بالضرورة أنَّ هذه المؤسسات رابحة، فقد تكون خاسرة في الواقع. أمّا ظهور النقصان في قيمة المؤشِّر فقد يعطي دلالات على الخسارة أو ضعف الإقبال على منتج هذه المؤسسات، فعلى سبيل المثال قد يُفسَّر النقصان لمؤشِّر المذياع على إحجام أكثرية الناس عن الاستماع إلى المذياع ومن ثَمّ قلة العروض التي تقدِّم إليها.

إذاً حتى الآن قمنا بتقديم المؤشِّر البسيط، ولكن قد يصادفنا في الكثير من الدراسات أرقاماً قياسيةً تنشأ من تغيُّرات القيم لسلع مختلفة. إنَّ المؤشِّر الموافق لمثل هذه الحالة يُدعى رقماً قياسياً مركّباً Composite Index Number، ومن هذه المؤشِّرات نقدِّم المؤشِّر الآتي.

### (٢ , ٢ , ١٣) تعريف (الرقم القياسي الكلي البسيط Simple Aggregate Index Number)

بفرض أنَّه لدينا $k$ شيء $T_1$، $T_2$،...و $T_k$ وكانت $p_{o,1}$ ، $p_{o,2}$ ،...و $p_{o,k}$ القيم الممثِّلة لهذه الأشياء في دورة الأساس على الترتيب، وكذلك $p_{c,1}$ ، $p_{c,2}$ ،...و $p_{c,k}$ القيم الممثِّلة لهذه الأشياء في دورة المقارنة على الترتيب أيضاً. عندئذ الرقم القياسي الكلي البسيط (سنرمز له بـ $\mathbf{I}_C$ ، ويُقرأ كنسبة مئوية) يُحسب بالعلاقة الآتية.

$$\mathbf{I}_C = \frac{\sum_{i=1}^{k} p_{c,i}}{\sum_{i=1}^{k} p_{o,i}} \times 100\% \qquad [13,3]$$

**(١٣,٢,٢,١) ملاحظة**

إنَّ الرقم القياسي الكلي البسيط يُوجد التغيُّر المئوي في تكلفة (أو سعر أو...) عدّة مواد من دورة إلى أخرى (أو من مكان إلى آخر)، فإذا كانت قيمة $\mathbf{I}_C$ أكبر من 100، فإنَّنا نطَرح 100 من قيمة هذا المؤشِّر، والنتيجة هي نسبة الزيادة في تكلفة (أو أسعار أو...) الأشياء من دورة الأساس إلى دورة المقارنة وتقرأ كنسبة مئوية. أمّا إذا كانت قيمة $\mathbf{I}_C$ أقل من 100، فعندئذ نطرح قيمة المؤشِّر من 100 والنتيجة هي النسبة المئوية لذلك النقص في تكلفة (أو أسعار أو...) الأشياء في دورة الأساس ممّا هي عليه في دورة المقارنة، وتقرأ كنسبة مئوية أيضاً.

**(١٣,٢,٢,٢) أمثلة**

١- بالعودة إلى المثال / ١/ من الأمثلة (١٣,٢,١,٢) (تكلفة المواد الغذائية المستهلكة من قبل أسرة)، فنجد أنَّ:

$$\sum_{i=1}^{4} p_{o,i} = 5+6+2+20 = 33 \qquad \& \qquad \sum_{i=1}^{4} p_{c,i} = 6+4+2+33 = 45$$

ومن ثمَّ (بحسب العلاقة [13,3]) يكون للرقم القياسي الكلي البسيط $\mathbf{I}_C$ لتكاليف المواد الغذائية المستهلكة القيمة الآتية:

$$\mathbf{I}_C = \frac{\sum_{i=1}^{4} p_{c,i}}{\sum_{i=1}^{4} p_{o,i}} \times 100 = \frac{45}{33} \times 100 = 136.36\%$$

وبما أنَّ قيمة الرقم القياسي الكلي البسيط $\mathbf{I}_C$ أكبر من 100 فإنَّ ذلك يُشير إلى زيادة عامّة في تكلفة المواد الغذائية التي قيد الدراسة من العام (1) إلى العام (2) بمقدار 36.36%.

٢- بالعودة إلى المثال / ٢/ من الأمثلة (١٣,٢,١,٢) (طلاب الثانوي الذين حصلوا على معدّل عام أعلى من 92%)، فنجد أنَّ:

$$\sum_{i=1}^{3} p_{o,i} = 0.65+0.52+0.83 = 2 \qquad \& \qquad \sum_{i=1}^{3} p_{c,i} = 0.75+0.60+0.78 = 2.13$$

ومن ثمَّ (بحسب العلاقة [13,3]) يكون لـ $\mathbf{I}_C$ الخاص بالطلاب الذين حصلوا على معدّل عام أعلى من 92% القيمة الآتية:

$$\mathbf{I}_C = \frac{2.13}{2} \times 100 = 106.5\%$$

وبما أنَّ قيمة $\mathbf{I}_C$ أكبر من 100 فإنَّ ذلك يشير إلى ارتفاع في نسبة طلاب الريف الذين حصلوا على معدل عام أعلى من 92% من أقرانهم من طلاب المدينة التي هي قيد الدراسة بمقدار 6.5% .

٣- بالعودة إلى المثال / ٣/ من الأمثلة (١٣,٢,١,٢) (تقوم شركة تجارية بحملة إعلامية سنوية لترويج منتجاتها)، فنجد أنَّ:

$$\sum_{i=1}^{4} p_{o,i} = 25+65+10+15 = 115 \qquad \& \qquad \sum_{i=1}^{4} p_{c,i} = 32+78+15+12 = 137$$

ومن ثمَّ (بحسب العلاقة [13,3]) يكون للرقم القياسي الكلي البسيط الخاص بتكاليف الإعلانات القيمة $\mathbf{I}_C = 119.13\%$، وبما أنَّ قيمة $\mathbf{I}_C$ أكبر من 100 فإنَّ ذلك يشير إلى زيادة عامّة في تكلفة الإعلانات من عام 2010 إلى عام 2014 بمقدار 19.13%.

**(١٣,٢,٢,٣) ملاحظة**

على الرغم من أن الرقم القياسي الكلي البسيط $\mathbf{I}_C$ سهل الحساب إلاّ أنّه يُعَاب عليه ما يلي:

١- إذا كان لأحد الأشياء وزناً (الذي يُمثّل بتكرار القيمة) كبيراً نسبياً، فعندئذ يمكن لهذا التكرار أن يسيطر على الرقم $\mathbf{I}_C$ ويجذبه إليه شأنه في ذلك كشأن القيم المتطرفة على المتوسط (الوسط الحسابي)، علماً أنَّ كلمة الوزن هنا لها المدلول نفسه المستخدم في الفصول السّابقة، فقد يعني الثقل لمادة ما، أو مضاعفات السعر لسلعة ما، أو التكرار لقيمة ما، أو...

٢- إذا كانت الأشياء مُدوّنة من أجل كميّات مختلفة فإنَّ الرقم القياسي الكلي البسيط $\mathbf{I}_C$ سيؤدي إلى أجوبة مختلفة، وهذا العيب يُعدّ من أسوأ العيوب للرقم $\mathbf{I}_C$؛ لأنَّ ذلك يجعل التلاعب في قيمة هذا المؤشّر ممكناً.

٣- إنّ الرقم $\mathbf{I}_C$ لا يأخذ في الحسبان كمية كل الأشياء المتداولة (أو السّلع المباعة).

## (١٣,٣) الأرقام القياسية الموزونة

## Weighted Index Numbers

لقد لاحظنا أنَّ الرقم القياسي الكلي البسيط $\mathbf{I}_C$ لا يأخذ بالحسبان تغيّر كميّات (أو أسعار أو تكلفة أو...) بعض الأشياء في أيّ من الدورتين الأساسية والمقارنة، ومن ثمَّ لا يمكن استخدامه إذا كان موضوع تغيّر كميّات (أو أسعار أو تكلفة أو...) بعض الأشياء مهمًّا بالنسبة إلى الدراسة المطلوبة. لذلك سنقدّم فيما يلي رقمين قياسيين من عائلة الأرقام القياسيّة الموزونة (أو المؤشّر الموزون). إنَّ هذين الرقمين القياسيين الموزونين يختلف بعضهما عن بعض في استخدام الأوزان، فالرقم الأول منهما (ويدعى مؤشّر لسبيريس **Laspeyres Index**) يستخدم أوزان دورة الأساس في حين أنَّ الرقم الآخر (ويدعى مؤشّر باش **Paasche Index**) يستخدم أوزان دورة المقارنة. إنَّ كلًّا منهما يقدّم حلاً جزئياً للمشكلة السّابقة ولكن من منطلق مختلف.

بالطبع إنَّ تعاملنا مع تغيّر قيم الأوزان (التكرارات أو الكميات أو...) $w_{o,1}$، $w_{o,2}$، ...و $w_{o,k}$ في دورة الأساس أو $w_{c,1}$، $w_{c,2}$، ...و $w_{c,k}$ في دورة المقارنة سيمنحنا مزيداً من المعلومات (بمعنى يصبح لدينا معلومات إضافية) حول هذه الأشياء، ومن ثمَّ سنحصل على نتائج أكثر مصداقيّة من الرقم $\mathbf{I}_C$ بخصوص الظاهرة التي هي قيد الدراسة.

فيما يلي نقدّم مؤشّرا يعود إلى الاقتصادي والإحصائي الألماني **لسبيريس (1834-1913)** Ernst Louis Étienne Laspeyres.

### (١٣,٣,١) تعريف (مؤشّر لسبيريس Laspeyres index)

بفرض أنّه لدينا $k$ شيء $T_1$، $T_2$، ...و $T_k$، وكانت $p_{o,1}$، $p_{o,2}$، ...و $p_{o,k}$ وكذلك $w_{o,1}$، $w_{o,2}$، ...و $w_{o,k}$ هي القيم الممثّلة لتلك الأشياء وأوزانها على الترتيب في دورة الأساس، وبالمثل $p_{c,1}$، $p_{c,2}$، ....و $p_{c,k}$ هي القيم الممثّلة لتلك الأشياء في دورة المقارنة على الترتيب. عندئذ تُدعى القيمة الآتية (والتي سنرمز لها بـ $\mathbf{I}_L$) مؤشّر لسبيريس لتلك الأشياء:

$$\mathbf{I}_L = \frac{\sum_{i=1}^{k} p_{c,i} \cdot w_{o,i}}{\sum_{i=1}^{k} p_{o,i} \cdot w_{o,i}} \times 100\ \% \qquad [13,4]$$

نلاحظ هنا أنَّ مؤشّر لسبيريس (وهو رقم قياسي موزون) يقيس التغيّر النسبي بناءً على تكرارات القيم في دورة الأساس فقط، وللتوضيح فلو كانت التكرارات هي كميّات السّلع مثلاً، فعندئذ مؤشّر لسبيريس يقيس التغيّر النسبي في تكلفة شراء هذه السّلع من الكميات المحدَّدة في دورة الأساس فقط.

#### (١٣,٣,١,١) أمثلة

١- بالعودة إلى المثال / ١/ من الأمثلة (١٣,٢,١,٢) حيث البيانات مقدَّمةً بالجدول (١٣,٢. أ)، وسنفترض أنَّ الكميات

المستهلكة من كل سلعة (مقدّرة بالكيلوغرام) في دورة الأساس كما في الجدول الآتي:

الجدول (٥ , ١٣ . أ)

| ر.م. $i$ | المنتج | الكميات المستهلكة في العام (1) Weights $w_{o,i}$ |
|---|---|---|
| 1 | الفاكهة | 400 |
| 2 | الخضراوات | 500 |
| 3 | الدقيق ومشتقاته | 200 |
| 4 | اللحوم | 300 |

ولنحسب مؤشِّر لسبيريس للمعطيات المقدَّمة آنفًا.

من أجل ذلك لنقم ببناء الجدول الآتي علماً أنَّ الكميّات في العام (1) مقدَّرة بال 100 كيلوغرام:

الجدول (٥ , ١٣ . ب)

| ر.م. $i$ | المنتج | الكميات في العام (1) $w_{o,i}$ | السعر في العام (1) Price – year (1) $p_{o,i}$ | السعر في العام (2) Price – year (2) $p_{c,i}$ | $p_{o,i} \cdot w_{o,i}$ | $p_{c,i} \cdot w_{o,i}$ |
|---|---|---|---|---|---|---|
| 1 | الفاكهة | 4 | 5 | 6 | 20 | 24 |
| 2 | الخضراوات | 5 | 6 | 4 | 30 | 20 |
| 3 | الدقيق | 2 | 2 | 2 | 4 | 4 |
| 4 | اللحوم | 3 | 20 | 33 | 60 | 99 |
| **Total** | ------------- | ------- | ------- | ------- | **114** | **147** |

ومن ثَمَّ باستخدام العلاقة [13,4] يكون لمؤشِّر لسبيريس القيمة $\mathbf{I}_L = \frac{147}{114} \times 100 = 128.95\ \%$، ومن ثَمَّ نجد وفقاً لمؤشِّر لسبيريس أنَّ الإنفاق قد ازداد بمقدار 28.95%.

٢- بالعودة إلى المثال / ٢/ من الأمثلة (٢ , ١ , ٢ , ١٣) (طلاب الثانوي الذين حصلوا على معدل عام أعلى من 92%)، وسنفترض أنَّه تمَّ اختيار عدد من المدن لكل مستوى من مستويات المرحلة الثانوية كما في الجدول الآتي:

الجدول (٦ , ١٣ . أ)

| ر.م. $i$ | المستوى | عدد المدن Weights $w_{o,i}$ |
|---|---|---|
| 1 | الأول الثانوي | 3 |
| 2 | الثاني الثانوي | 2 |
| 3 | الثالث الثانوي | 5 |

فعندئذ لحساب مؤشِّر لسبيريس لتلك المعطيات سنقوم ببناء الجدول الآتي:

الجدول (٦، ١٣. ب)

| ر.م. $i$ | المستوى | عدد المدن $w_{o,i}$ | النسب في المدن $p_{o,i}$ | النسب في الأرياف $p_{c,i}$ | $p_{o,i} \cdot w_{o,i}$ | $p_{c,i} \cdot w_{o,i}$ |
|---|---|---|---|---|---|---|
| 1 | الأول الثانوي | 3 | 0.65 | 0.75 | 1.95 | 2.25 |
| 2 | الثاني الثانوي | 2 | 0.52 | 0.60 | 1.04 | 1.2 |
| 3 | الثالث الثانوي | 5 | 0.83 | 0.78 | 4.15 | 3.9 |
| Total | ------------- | ------- | ------- | ------- | 7.14 | 7.35 |

ومن ثَمّ باستخدام العلاقة [13,4] يكون لمؤشّر لسبيريس القيمة $\mathbf{I}_L = \dfrac{7.35}{7.14} \times 100 = 102.94\%$، ومن ثَمَّ نجد وفقاً لمؤشّر لسبيريس أنَّه ما زال هناك ارتفاع بمقدار 2.94% في نسبة طلاب الثانوي الذين حصلوا على معدل عام أعلى من 92% لدى أبناء الريف على أقرانهم من أبناء المدن.

٣- بالعودة إلى المثال / ٢/ من الأمثلة (٢، ١، ٢، ١٣) (تقوم شركة تجارية بحملة إعلامية سنوية لترويج منتجاتها)، وسنفترض أنَّه تمَّ التعاقد مع عدد من المؤسسات الإعلامية في عام 2010 وفقاً للجدول الآتي:

الجدول (٧، ١٣. أ)

| ر.م. $i$ | المؤسسة الإعلامية في عام 2014 | عدد مؤسسات Weights $w_{o,i}$ |
|---|---|---|
| 1 | الصحافة | 3 |
| 2 | الرائي | 4 |
| 3 | إعلانات الطرق | 2 |
| 4 | المذياع | 2 |

فعندئذ لحساب مؤشّر لسبيريس لتلك المعطيات نقوم ببناء الجدول الآتي:

الجدول (٧، ١٣. ب)

| ر.م. $i$ | المؤسسة الإعلامية | عدد المؤسسات $w_{o,i}$ | مدفوعات المؤسسة في العام 2010 $p_{o,i}$ | مدفوعات المؤسسة في العام 2014 $p_{c,i}$ | $p_{o,i} \cdot w_{o,i}$ | $p_{c,i} \cdot w_{o,i}$ |
|---|---|---|---|---|---|---|
| 1 | الصحافة | 3 | 25 | 32 | 75 | 96 |
| 2 | الرائي | 4 | 65 | 78 | 260 | 312 |
| 3 | إعلانات الطرق | 2 | 10 | 15 | 20 | 30 |
| 4 | المذياع | 2 | 15 | 12 | 30 | 24 |
| Total | ------------- | ------- | ------- | ------- | 385 | 462 |

وباستخدام العلاقة [13,4] يكون لمؤشّر لسبيريس القيمة $\mathbf{I}_L = \frac{462}{385} \times 100 = 120$، ومن ثمّ نجد وفقاً لمؤشّر لسبيريس أنَّ الزيادة في التكاليف قد بلغت 20%.

لقد لاحظنا أنَّ مؤشّر لسبيريس يهتمّ بكميّة المواد في دورة الأساس فقط، فإذا كانت الدراسة تتعلّق بالكميّات في دورة المقارنة فإنَّ ذلك المؤشّر لا يفيدنا بشيء، ولذلك سنقدّم فيما يلي مؤشّراً يأخذ في الحسبان كميّات المواد في دورة المقارنة، ولكنّه لا يهتمّ بكميات دورة الأساس. إنَّ هذا المؤشّر يقدّمه لنا التعريف الآتي الذي يعود إلى الاقتصادي والإحصائي الألماني باشّي Hermann Paasche (1851-1925).

### (١٣,٣,٢) تعريف (مؤشّر باشّي Paasche index)

بفرض أنّه لدينا $k$ شيء $T_1$، $T_2$،...و $T_k$، وكانت $p_{o,1}$ ، $p_{o,2}$ ،...و $p_{o,k}$ القيم الممثّلة لهذه الأشياء في دورة الأساس على الترتيب، وكذلك سنفترض أنَّ $p_{c,1}$ ، $p_{c,2}$ ،....و $p_{c,k}$ و $w_{c,1}$، $w_{c,2}$،...و $w_{c,k}$ القيم الممثّلة لهذه الأشياء وأوزانها على الترتيب في دورة المقارنة. عندئذ تُحسب قيمة الرقم القياسي الموزون لـ باشّي (الذي سنرمز له بـ $\mathbf{I}_P$) باستخدام العلاقة الآتية:

$$\mathbf{I}_P = \frac{\sum_{i=1}^{k} p_{c,i} \cdot w_{c,i}}{\sum_{i=1}^{k} p_{o,i} \cdot w_{c,i}} \times 100\ \% \qquad [13,5]$$

نلاحظ هنا أنَّ مؤشّر باشّي (وهو رقم قياسي موزون أيضاً) يقيس التغيّر النسبي بناءً على تكرارات القيم في دورة المقارنة فقط، وللتوضيح فلو كانت التكرارات هي كميّات السّلع مثلاً فعندئذ مؤشّر باشّي يقيس التغيّر النسبي في تكلفة شراء هذه السّلع من الكميّات المُحدَّدة في دورة المقارنة فقط.

#### (١٣,٣,٢,١) أمثلة

١- بالعودة إلى المثال /١/ من الأمثلة (١٣,٢,١,٢) (المواد الغذائية المستهلكة من قبل أسرة)، وسنفترض أنَّ الكميات المستهلكة (مقدَّرة ب 100 الكيلو غرام) من السّلع المذكورة في العام الآخر (العام (2) أو دورة المقارنة) مقدَّمة كما في الجدول الآتي علماً أنَّ الكميّات في العام (2) مقدَّرة بال 100 كيلوغرام:

الجدول (١٣,٨.أ)

| ر.م. $i$ | المنتج | الكميات المستهلكة في العام (2) Weights $w_{c,i}$ |
|---|---|---|
| 1 | الفاكهة | 6 |
| 2 | الخضراوات | 8 |
| 3 | الدقيق ومشتقاته | 3 |
| 4 | اللحوم | 5 |

فعندئذ لحساب مؤشّر باشّي $\mathbf{I}_P$ سنقوم ببناء الجدول الآتي:

الجدول (٨, ١٣. ب)

| ر.م. $i$ | المنتج | الكميات في العام (2) $w_{c,i}$ | السعر في العام (1) Price – year (1) $p_{o,i}$ | السعر في العام (2) Price – year (2) $p_{c,i}$ | $p_{o,i} \cdot w_{c,i}$ | $p_{c,i} \cdot w_{c,i}$ |
|---|---|---|---|---|---|---|
| 1 | الفاكهة | 6 | 5 | 6 | 30 | 36 |
| 2 | الخضراوات | 8 | 6 | 4 | 48 | 32 |
| 3 | الدقيق ومشتقاته | 3 | 2 | 2 | 6 | 6 |
| 4 | اللحوم | 5 | 20 | 33 | 100 | 165 |
| Total | ------------- | ------- | ------- | ------- | 184 | 239 |

ومنه باستخدام العلاقة [13,5] يكون لمؤشّر باشّي القيمة $\mathbf{I}_P = \frac{239}{184} \times 100 = 129.89$، فلاحظ هنا أنَّ قيمة مؤشّر لسبيريس تأخذ قيمة أقل من مؤشّر باشّي، وهذا يدلّ على أنَّ استخدام مؤشّر لسبيريس يتزايد بشكل أسرع في قطاعات السعر الأعلى.

٢- بالعودة إلى المثال / ٢/ من الأمثلة (٢, ١, ٢, ١٣) (طلاب الثانوي الذين حصلوا على معدّل عام أعلى من 92%)، وسنفترض مع ثبات النسب السابقة لطلاب المرحلة الثانوي الذين حصلوا على معدل عام أعلى من 92% أنّه قد تمَّ اختيار عشوائي لبعض الأرياف من كل مدينة موزّعة على النحو المقدَّم في الجدول الآتي:

الجدول (٩, ١٣. أ)

| ر.م. $i$ | المستوى | عدد الأرياف من المدينة Weights $w_{c,i}$ |
|---|---|---|
| 1 | الأول الثانوي | 4 |
| 2 | الثاني الثانوي | 3 |
| 3 | الثالث الثانوي | 5 |

فعندئذ لحساب مؤشّر باشّي لهذه المعطيات سنبني الجدول الآتي:

جدول (٩, ١٣ ب).

| ر.م. $i$ | المستوى | عدد الأرياف من المدينة $w_{c,i}$ | النسب في المدن $p_{o,i}$ | النسب في الأرياف $p_{c,i}$ | $p_{o,i} \cdot w_{c,i}$ | $p_{c,i} \cdot w_{c,i}$ |
|---|---|---|---|---|---|---|
| 1 | الأول الثانوي | 4 | 0.65 | 0.75 | 2.6 | 3 |
| 2 | الثاني الثانوي | 3 | 0.52 | 0.60 | 1.56 | 1.8 |
| 3 | الثالث الثانوي | 5 | 0.83 | 0.78 | 4.15 | 3.9 |
| Total | ------------- | ------- | ------- | ------- | 8.31 | 8.7 |

ومن ثمّ باستخدام العلاقة [13,5] يكون لمؤشّر باشّي القيمة $\mathbf{I}_P = \frac{8.7}{8.31} \times 100 = 104.69\%$ ، ومن ثمَّ نجد وفقاً لمؤشّر باشّي أنّه ما زال هناك ارتفاع بمقدار 4.69% في نسبة طلاب الثانوي الذين حصلوا على معدل عام أعلى من 92% لدى أبناء الريف على أقرانهم من أبناء المدن.

٣- بالعودة إلى المثال / ٣/ من الأمثلة (٢, ١, ٢, ١٣) (تقوم شركة تجارية بحملة إعلامية سنوية لترويج منتجاتها)، وسنفترض أنَّه تمَّ التعاقد مع عدد من المؤسسات الإعلامية في عام 2014 وفقاً للجدول الآتي:

الجدول (١٠, ١٣. أ)

| ر.م. $i$ | المؤسسة الإعلامية في عام 2014 | عدد مؤسسات Weights $w_{c,i}$ |
|---|---|---|
| 1 | الصحافة | 4 |
| 2 | الرائي | 12 |
| 3 | إعلانات الطرق | 3 |
| 4 | المذياع | 2 |

فعندئذ لحساب مؤشّر باشّي لتلك المعطيات نقوم بإنشاء الجدول الآتي:

الجدول (١٠, ١٣. ب)

| ر.م. $i$ | المؤسسة الإعلامية | عدد المؤسسات $w_{c,i}$ | مدفوعات المؤسسة في عام 2010 $p_{o,i}$ | مدفوعات المؤسسة في عام 2014 $p_{c,i}$ | $p_{o,i} \cdot w_{c,i}$ | $p_{c,i} \cdot w_{c,i}$ |
|---|---|---|---|---|---|---|
| 1 | الصحافة | 4 | 25 | 32 | 100 | 128 |
| 2 | الرائي | 12 | 65 | 78 | 780 | 936 |
| 3 | إعلانات الطرق | 3 | 10 | 15 | 30 | 45 |
| 4 | المذياع | 2 | 15 | 12 | 30 | 24 |
| Total | ------------- | ------- | ------- | ------- | 940 | 1133 |

ومنه باستخدام العلاقة [13,5] يكون لمؤشّر باشّي القيمة $\mathbf{I}_P = \frac{1133}{940} \times 100 = 120.5\ \%$ ، ومن ثمَّ نجد وفقاً لمؤشّر باشّي أنَّ الزيادة في التكاليف قد بلغت 20.5%.

### (٢, ٢, ٣, ١٣) مقارنة بين مؤشّري لسبيريس و باشّي

١ – من وجهة نظر الاقتصاد يُلاحظ أنَّ مؤشّر لسبيريس يقيس نسبة الإنفاق على كميّات في دورة الأساس وفقاً لتكلفة دورة المقارنة إلى الإنفاق على تلك الكميّات وفقاً لتكلفة دورة الأساس، بينما مؤشّر باشّي يقيس نسبة الإنفاق على كميّات في دورة المقارنة وفقاً لتكلفة دورة المقارنة إلى الإنفاق على تلك الكميّات وفقاً لتكلفة دورة الأساس.

٢- بما أنَّ مؤشّر لسبيريس يستخدم أوزان دورة الأساس، فإنّه من وجهة نظر الاقتصاد يُلاحظ أنَّ مؤشّر لسبيريس قد يبالغ في تقدير الارتفاع في تكاليف المعيشة (لأن النّاس قد خفضت استهلاكها من العناصر التي أصبحت أغلى نسبيا من غيرها)، في حين أنَّ مؤشّر باشّي يستخدم أوزان دورة المقارنة، ولذلك فقد يظهر أنّه يقلّل من الارتفاع في تكاليف المعيشة.

٣- قيمة مؤشّر لسبيريس تكون عادةً أكبر من قيمة مؤشّر باشّي، أي إنَّ $\mathbf{I}_L \geq \mathbf{I}_P$ .

٤- يتطلب مؤشّر باشّي مجموعة جديدة من الأوزان التي يجب الحصول عليها في كل عام، وهذه المعلومات قد يكون الحصول عليها مكلفاً أو شاقاً، ولذلك يصعب استخدام مؤشّر باشّي لإجراء مقارنات من سنة إلى أخرى.

٥- بالرغم من أنّه لكلٍّ من مؤشّري لسبيريس وباشّي مهمّة مُحَدَّدة، ولا يغني أحدهما عن الآخر في كثير من الحالات. إلاَّ أنّه بسبب الفقرتين الأخيرتين / ٥ و٦/ يرجَّح استخدام مؤشّر لسبيريس على مؤشّر باشّي في كثير من الحالات.

إذاً، وكما لاحظنا أنَّ كلًّا من مؤشّري لسبيريس وباشّي يقوم بمهمّة محدَّدة وبحيث لا يغني أحدهما عن الآخر، فإذا كانت الدراسة مركّزة على ما يتمّ استهلاكه في دورة الأساس ودورة المقارنة أيضاً، فعندئذ كلّ من المؤشّرين السّابقين يصبح قليل الفائدة، ولهذا سنقدِّم فيما يلي مؤشّراً يأخذ في الحسبان كميّات المواد في دورة الأساس ودورة المقارنة معاً. إنَّ هذا المؤشّر يُعرف باسم **المؤشّر المثالي لـ فيشر** ويقدّمه لنا التعريف الآتي.

### (١٣,٣,٣) تعريف (المؤشّر المثالي لـ فيشر Fisher's ideal index)

بفرض أنّه لدينا $k$ شيء $T_1$، $T_2$،...و $T_k$، وكانت $p_{o,1}$ ، $p_{o,2}$ ،...و $p_{o,k}$ وكذلك $w_{o,1}$، $w_{o,2}$،...و $w_{o,k}$ القيم الممثّلة لهذه الأشياء وأوزانها على الترتيب في دورة الأساس، وكذلك $p_{c,1}$ ، $p_{c,2}$ ،....و $p_{c,k}$ وكذلك $w_{c,1}$، $w_{c,2}$،...و $w_{c,k}$ القيم الممثّلة لهذه الأشياء وأوزانها على الترتيب في دورة المقارنة. عندئذ القيمة الآتية (والتي سنرمز لها بـ $\mathbf{I}_F$ ) تُدعى **المؤشّر المثالي لـ فيشر**:

$$\mathbf{I}_F = \sqrt{\mathbf{I}_L \cdot \mathbf{I}_P}\ \% \qquad [13,6]$$

علماً أنَّ $\mathbf{I}_L$ و $\mathbf{I}_P$ هما قيمة مؤشّري لسبيريس وباشي على الترتيب للمعطيات السابقة.

لاحظ هنا أنَّ المؤشر المثالي لـ فيشر متعلّق بمؤشري لسبيريس وباش (المتوسط الهندسي لهما)، ولذلك فإنَّ هذا المؤشّر يقيس التغيّر النسبي بناءً على تكرارات قيم دورة الأساس والمقارنة معاً، وللتوضيح فلو كانت التكرارات هي كميّات السّلع مثلاً، فعندئذ المؤشّر المثالي لـ فيشر يقيس التغيّر النسبي في تكلفة شراء هذه السّلع من الكميات المحدَّدة في دورتي الأساس والمقارنة بآنٍ واحد.

### (١٣,٣,٣,١) أمثلة

١- بالعودة إلى المثال / ١/ من الأمثلة (١٣,٢,١,٢) (المواد الغذائية المستهلكة من قبل أسرة)، سنفترض علاوةً على ما سبق أنَّ الكميات المستهلكة (مقدّرة بـ 100 الكيلو غرام) من السّلع المذكورة في العام الآخر (العام (2) أو دورة المقارنة) مقدَّمة كما في الجدول (١٣,٨ . أ) حيث وجدنا سابقاً أنَّ لمؤشّر لسبيريس القيمة $\mathbf{I}_L = 128.95\ \%$ ، وكذلك لمؤشّر باشّي القيمة $\mathbf{I}_P = 129.89\ \%$ ، ومن ثَمّ يكون للمؤشّر المثالي لـ فيشر القيمة الآتية:

$$\mathbf{I}_F = \sqrt{\mathbf{I}_L \cdot \mathbf{I}_P} = \sqrt{128.95 \times 129.89} = 129.42\,\%$$

ومن ثَمَّ نجد وفقاً للمؤشّر المثالي لـ فيشر أنَّ الزيادة في أسعار السّلع قد بلغت 29.42%.

٢- بالعودة إلى المثال / ٢/ من الأمثلة (١٣,٢,١,٢) (**طلاب الثانوي الذين حصلوا على معدل عام أعلى من 92%**)، سنفترض علاوةً على ما سبق أنَّ نسب طلاب المرحلة الثانوي الذين حصلوا على معدل عام أعلى من 92% مقدَّمة كما في الجدول (١٣,٩. أ) حيث وجدنا سابقاً أنَّ لمؤشِّر لسبيريس القيمة $\mathbf{I}_L = 102.94\ \%$، وكذلك لمؤشِّر باشّي القيمة $\mathbf{I}_P = 104.69\ \%$، ومنه يكون للمؤشِّر المثالي لـ فيشر القيمة الآتية:

$$\mathbf{I}_F = \sqrt{\mathbf{I}_L \cdot \mathbf{I}_P} = \sqrt{102.94 \times 104.69} = 103.81\ \%$$

ومن ثَمَّ نجد وفقاً للمؤشِّر المثالي لـ فيشر أنَّ الزيادة في نسب طلاب المرحلة الثانوي الذين حصلوا على معدّل عام أعلى من 92% في الأرياف عن أقرانهم في المدن قد بلغت 3.81%.

٢- بالعودة إلى المثال / ٣/ من الأمثلة (١٣,٢,١,٢) (**تقوم شركة تجارية بحملة إعلامية سنوية لترويج منتجاتها**)، سنفترض علاوةً على ما سبق أنَّ عدد المؤسسات الإعلامية المتعاقد معها في عام 2014 مقدَّمة كما في الجدول (١٣,١٠. أ) حيث وجدنا سابقاً أنَّ لمؤشِّر لسبيريس القيمة $\mathbf{I}_L = 120\ \%$، وكذلك لمؤشِّر باشّي القيمة $\mathbf{I}_P = 120.5\ \%$، ومن ثَمّ يكون للمؤشِّر المثالي لـ فيشر القيمة الآتية:

$$\mathbf{I}_F = \sqrt{\mathbf{I}_L \cdot \mathbf{I}_P} = \sqrt{120 \times 120.5} = 120.25\ \%$$

ومن ثَمَّ نجد وفقاً للمؤشِّر المثالي لـ فيشر أنَّ الزيادة في تكاليف الحملة الدعائية قد بلغت 20.25%.

## (١٣,٤) أمثلة
### Examples

١- لنفترض أنَّ إحدى المؤسسات العلمية المهتمّة بالبيئة تقوم بدراسة التحولات التي تطرأ على كوكب الأرض، واعتمدت من أجل ذلك معايير معيَّنة تحدِّد من خلالها حالة المنطقة التي خضعت للدراسة أهي: ممتازة، جيدةً جداً، جيدة، مقبولة، سيئة، أم سيئة جداً، وأنَّ هذه المؤسسة قد قامت بهذه الدراسة على أربع مناطق متمايزة من كوكب الأرض هي: مناطق الغابات، والبوادي، والمرتفعات الجبلية، والمناطق القطبية. كذلك سنفترض أنَّها قامت بهذه الدراسة في فترتين زمنيتين قبل 50 و25 سنة وهذا العام أيضاً، وأنَّ النتائج لهذه الدراسة كانت كما في الجدول الآتي:

الجدول (١٣,١١. أ)

| منطقة | حالة المنطقة الآن | حالة المنطقة قبل 25 سنة | حالة المنطقة قبل 50 سنة |
|---|---|---|---|
| غابات | مقبولة | جيدة | جيدة جداً |
| البوادي | سيئة جداً | مقبولة | جيدة |
| مرتفعات جبلية | مقبولة | جيدة | جيدةً جداً |
| قطبية | سيئة | جيدة | ممتازة |

ولننظر ما آلت إليه هذه المناطق على ما كانت عليه قبل 25 سنة مستخدمين في ذلك الأرقام لقياسيّة.

من أجل هذه الدراسة سوف نستخدم الأرقام 6 و5 و4 و3 و2 و1 للتعبير عن العبارات: ممتازة، جيدةً جداً، جيدة، مقبولة،

سيئة، سيئة جداً على الترتيب. كذلك يُلاحظ أنَّ حالة المناطق قبل 25 سنة ستعتمد كدورة أساس في حين يصبح لدينا دورتي مقارنة هما حالة المناطق قبل 50 سنة، وكذلك حالة المناطق في العام الحالي، ومن ثمَّ يمكننا أن نعرض الجدول الآتي لهذه المعطيات:

الجدول (١١ ,١٣ . ب)

| دورة مقارنة<br>حالة المنطقة قبل 50 سنة | دورة الأساس<br>حالة المنطقة قبل 25 سنة | دورة مقارنة<br>حالة المنطقة الآن | المنطقة | ر.م.<br>$i$ |
|---|---|---|---|---|
| 5 | 4 | 3 | غابات | 1 |
| 4 | 3 | 2 | البوادي | 2 |
| 5 | 4 | 3 | مرتفعات جبلية | 3 |
| 6 | 4 | 2 | قطبية | 4 |
| 20 | 15 | 10 | ------------ | Total |

هذا من جانب ، ومن جانب آخر يُلاحظ أنَّ الأرقام القياسيّة التي يُمكننا استخدمها هنا هي الأرقام القياسيّة البسيطة فقط (بسبب عدم وجود تكرارات – أوزان –للقيم الممثّلة للحالات).

الآن لو أخذنا الفترة قبل 50 سنة كدورة مقارنة فإنَّنا نجد من العلاقة [13,2] أنَّ:

أ- قيمة المؤشّر البسيط لمنطقة الغابات تساوي:

$$\mathbf{I}_{S,1} = \frac{p_{c,1}}{p_{o,1}} \times 100 = \frac{5}{4} \times 100 = 125 \ \%$$

وهذا يعني أنَّ حالة الغابات قبل 50 سنة كان أفضل بـ %25 ممّا كانت عليه قبل 25 سنة، بمعنى أنَّ حالة الغابات قبل 25 سنة قد تردت بنسبة %25 أكثر ممّا كانت عليه قبل 50 سنة. في الحقيقية نريد أن نبيّن من خلال هذا الحوار أنَّ ازدياد النسبة عن %100 لا يعني بالضرورة أنَّ شيئاً حسناً قد حصل، وكذلك لتوضيح أنَّ القيم التي نحصل عليها في مثل هذه الدراسات تفهم بشكل نسبي، وتفسيرها يتبع هذه النسبية.

ب- قيمة المؤشّر البسيط لمنطقة البوادي تساوي $\mathbf{I}_{S,2} = \frac{4}{3} \times 100 = 133.33 \ \%$، وهنا نلاحظ أنَّ حالة البوادي قبل 25 سنة قد تردت بنسبة %33.33 أكثر ممّا كانت عليه قبل 50 سنة.

ج- منطقة المرتفعات الجبلية تساوي $\mathbf{I}_{S,3} = 125 \ \%$، وهنا نلاحظ أنَّ حالة المرتفعات الجبلية قبل 25 سنة قد تردت بنسبة %25 أكثر ممّا كانت عليه قبل 50 سنة.

د- المنطقة القطبية تساوي $\mathbf{I}_{S,4} = 150 \ \%$، وهنا نلاحظ أنَّ حالة القطبية قبل 25 سنة قد تردت بنسبة %50 أكثر ممّا كانت عليه قبل 50 سنة.

هكذا نجد أنَّ قيمة معدَّل المؤشّر البسيط لحالات هذه المناطق تساوي:

$$\mathbf{SA} := \frac{1}{4}\sum_{i=1}^{4} \ \mathbf{I}_{S,i} = \frac{125+133.33+125+150}{4} = \frac{533.33}{4} = 133.333 \ \%$$

فنلاحظ أنَّ حالة المناطق التي قيد الدراسة قبل 25 سنة قد تردت بمعدَّل %33.33 أكثر ممّا كانت عليه قبل 50 سنة.

أمّا المؤشّر الكلي البسيط لحالات هذه المناطق فنجده باستخدام العلاقة [13,3] يساوي:

$$\mathbf{I}_C = \frac{20}{15} \times 100 = 133.333\ \%$$

وهي ذات القيمة التي حصلنا عليها من أجل قيمة معدَّل المؤشِّر البسيط لحالات هذه المناطق. لكن يجب الانتباه هنا إلى أنَّه **ليس بالضرورة** أن تكون قيمة معدَّل المؤشِّر البسيط تساوي قيمة المؤشِّر الكلي البسيط.

بشكل مماثل يمكن للقارئ تنفيذ الحسابات السابقة عندما نأخذ حالة المناطق في العام الحالي كدورة مقارنة.

الآن، لنقم بتمديد دراستنا السابقة، وسنفترض أنَّه خضعت مواقع عديدة من كل منطقة من المناطق التي هي قيد الدراسة، وأنَّ النتائج كانت كما في الجدول الآتي:

الجدول (١١ ,١٣ .ج)

| ر.م. $i$ | المنطقة | دورة الأساس قبل 25 سنة | | دورة المقارنة قبل 50 سنة | |
|---|---|---|---|---|---|
| | | عدد المواقع | حالة المنطقة قبل 25 سنة | عدد المواقع | حالة المنطقة قبل 50 سنة |
| 1 | غابات | 5 | 4 | 6 | 5 |
| 2 | البوادي | 4 | 3 | 3 | 4 |
| 3 | مرتفعات جبلية | 3 | 4 | 5 | 5 |
| 4 | قطبية | 5 | 4 | 4 | 6 |
| Total | ------------- | 17 | 15 | 18 | 20 |

ولنقم بحساب قيمة المؤشِّر المثالي لـ فيشر $\mathbf{I}_F$ للمعطيات السابقة.

من أجل ذلك علينا حساب مؤشِّري لسبيريس $\mathbf{I}_L$ وباشّي $\mathbf{I}_P$، فنجد باستخدام العلاقتين [13,4] و[13,5] أنَّ:

$$\mathbf{I}_L = \frac{86}{64} \times 100 = 134.38\ \% \qquad \& \qquad \mathbf{I}_P = \frac{91}{69} \times 100 = 149.18\ \%$$

ومنه باستخدام العلاقة [13,6] نجد أنَّ قيمة المؤشِّر المثالي لـ فيشر تساوي:

$$\mathbf{I}_F = \sqrt{\mathbf{I}_L \cdot \mathbf{I}_P} = \sqrt{134.38 \times 149.18} = 141.59\ \%$$

وهذا يشير إلى أنَّ حالة المواقع التي خضعت للدراسة قبل 25 سنة قد تردت بنسبة 41.59% أكثر ممّا كانت عليه قبل 50 سنة وفقاً للمواقع التي خضعت للدراسة آنذاك.

بشكل مماثل يمكن للقارئ تنفيذ الحسابات السابقة عندما نأخذ حالة المناطق في العام الحالي كدورة مقارنة.

٢- لنفترض أنَّ حكومة دولة ما أرادت تقدِّيم زيادة عامة لرواتب الموظفين في الدولة، ولكي تكون الزيادة منصفة رأت أنَّ قيمة المؤشِّر المثالي لـ فيشر يعطيها تبصراً جيداً لنسبة الزيادة المقترحة، فقامت بسبر لأسعار متطلبات الحياة المعيشية للمواطنين (مقدَّراً بـ 100 وحدة نقدية) ولكل متطلب على حدة في عدد من المدن (اختيرت عشوائياً) عند آخر زيادة منحت للموظفين، ومن ثمَّ قامت من أجل كل متطلب على حدة بالنظر في أسعار تلك المتطلبات في العام الحالي لعدد من المدن (اختيرت عشوائياً أيضاً)، فكانت لديها المُعطيات الآتية:

الجدول (١٢, ١٣. أ)

| ر.م. $i$ | المتطلب | دورة الأساس العام الذي تمت فيه آخر زيادة للراتب | | دورة المقارنة العام الحالي | |
|---|---|---|---|---|---|
| | | عدد المدن الخاضعة للدراسة | متوسط التكلفة الشهري | عدد المدن الخاضعة للدراسة | متوسط التكلفة الشهري |
| 1 | المواد الغذائية | 8 | 12 | 7 | 20 |
| 2 | الطاقة | 5 | 5 | 5 | 7 |
| 3 | وسائل النقل | 7 | 8 | 6 | 10 |
| 4 | الطبابة | 5 | 7 | 8 | 8 |
| Total | ------------- | 25 | 32 | 26 | 45 |

الآن لحساب قيمة المؤشِّر المثالي لـ فيشر $\mathbf{I}_F$ للمعطيات السابقة علينا حساب مؤشِّري لسبيريس $\mathbf{I}_L$ و باشّي $\mathbf{I}_P$، فنجد باستخدام العلاقتين [13,4] و[13,5] أنَّ:

$$\mathbf{I}_L = \frac{305}{212} \times 100 = 143.87\ \% \qquad \& \qquad \mathbf{I}_P = \frac{299}{213} \times 100 = 140.38\ \%\ \ \%$$

ومنه باستخدام العلاقة [13,6] نجد أنَّ قيمة المؤشِّر المثالي لـ فيشر تساوي:

$$\mathbf{I}_F = \sqrt{\mathbf{I}_L \cdot \mathbf{I}_P} = \sqrt{143.87 \times 140.38} = 142.11\ \%$$

وهذا يشير إلى أنَّ متطلبات الحياة المعيشية للمواطنين في هذا العام قد ازدادت بنسبة 42.11% أكثر ممّا كانت عليه لدى آخر زيادة للراتب، ولذلك من الممكن أن تقترح زيادة على الرواتب الشهرية للموظفين بنسبة 42.11% من الراتب الأساسي.

لقد ذكرنا سابقاً أنَّه يجب التفكير مليًّا عند اختيار دورة الأساس وبشكل خاص عند استخدام المؤشِّرات الموزونة، فعلى سبيل المثال لو قمنا بعكس الفرضيات السابقة المتعلِّقة بدورة الأساس فعندئذ سيكون لدينا المُعطيات الآتي:

الجدول (١٢, ١٣. ب)

| ر.م. $i$ | المتطلب | دورة الأساس العام الحالي | | دورة مقارنة العام الذي تمت فيه آخر زيادة للراتب | |
|---|---|---|---|---|---|
| | | عدد المدن الخاضعة للدراسة | متوسط التكلفة الشهري | عدد المدن الخاضعة للدراسة | متوسط التكلفة الشهري |
| 1 | المواد الغذائية | 7 | 20 | 8 | 12 |
| 2 | الطاقة | 5 | 7 | 5 | 5 |
| 3 | وسائل النقل | 6 | 10 | 7 | 8 |
| 4 | الطبابة | 8 | 8 | 5 | 7 |
| Total | ------------- | 26 | 45 | 25 | 32 |

ومن ثمَّ يكون لدينا:

$$\mathbf{I}_L = \frac{213}{299}\times 100 = 71.24\ \% \qquad \& \qquad \mathbf{I}_P = \frac{212}{305}\times 100 = 69.51\ \%$$

ومنه باستخدام العلاقة [13,6] نجد أنَّ قيمة المؤشِّر المثالي لـ فيشر تساوي:

$$\mathbf{I}_F = \sqrt{\mathbf{I}_L \cdot \mathbf{I}_P} = \sqrt{71.24\times 69.51} = 70.37\ \%$$

وهذا يشير إلى أنَّ متطلبات الحياة المعيشية للمواطنين لدى آخر زيادة للراتب كانت أقل بنسبة 29.76% ممّا هي عليه في العام الحالي، وهذه القيمة تختلف كثيراً عن القيمة التي حصلنا عليها سابقاً.

**هذا ما تيسَّر تقديمه من أجل الجزء النظري والتطبيقي لهذا الفصل**

# تمارين الفصل الثالث عشر

١- اشترك أربعة أشخاص كل على حدة في البورصة حيث اشترى كل منهم سهماً من شركة ما بمبلغ (مقدّر بوحدة نقدية ما)، ومن ثمَّ باعه لاحقاً بقيمة كما في الجدول الآتي:

| ر.م. $i$ | الشخص | سعر الشراء | سعر البيع |
|---|---|---|---|
| 1 | A | 1500 | 1600 |
| 2 | B | 1600 | 1550 |
| 3 | C | 1450 | 1600 |
| 4 | D | 1550 | 1500 |

والمطلوب ما يلي:

أ- عيّن دورتي الأساس والمقارنة مع تعليل السبب إن كان هناك تفضيل لأحدهما على الآخر.

ب- احسب المؤشِّرات البسيطة للمُعطيات السابقة.

ج- احسب المؤشِّر الكلي البسيط للمُعطيات السابقة، وماذا تستنتج من ذلك؟

٢- تقدَّم ثلاثة متسابقين إلى برنامج مسابقات رياضية مكوَّن من مرحلتين، وبحيث يمكن للمتسابق أن يحصل على درجة أقصاها 100%، وعلاوة على ذلك يسمح للمتسابق في المرحلة الأولى أن يجري حتى خمسة تدريبات قبل دخول المسابقة، وبعد الانتهاء من المرحلة الأُولى يمكنه دخول المرحلة الثانية دون أيّة تدريبات مُسبقة. الآن سنفترض أنَّ النتائج كانت كما في الجدول الآتي:

| المتسابق | عدد التدريبات السابقة للمرحلة الأولى | الدرجة التي حصل عليها المتدرب في المرحلة الأولى | الدرجة التي حصل عليها المتدرب في المرحلة الثانية |
|---|---|---|---|
| A | 4 | 85% | 75% |
| B | 3 | 92% | 95% |
| C | 5 | 87% | 82% |

والمطلوب ما يلي:

أ - عيّن دورتي الأساس والمقارنة مع تعليل السبب إن كان هناك تفضيل لأحدهما على الآخر.

ب- احسب المؤشِّرات البسيطة للمُعطيات السابقة.

ج- احسب المؤشِّر الكلي البسيط للمُعطيات السابقة، وماذا تستنتج من ذلك؟

د- إذا اخترنا المرحلة الأولى كدورة أساس، فأي المؤشرات الموزونة صالح للاستخدام في هذه الحالة، ومن ثمَّ احسب قيمة هذا المؤشِّر؟

هـ- إذا اخترنا المرحلة الثانية كدورة أساس، فأي المؤشرات الموزونة صالح للاستخدام في هذه الحالة، ومن ثمَّ احسب قيمة هذا المؤشِّر؟

٣- يقوم بلد باستيراد بعض المواد والسلع وكذلك يصدِّر من منتجاته مواد وسلعًا مماثلة (كأن يستورد القمح الذي يستخدم لصناعة الخبز ويصدِّر القمح الذي يستخدم لصناعة البسكويت)، وسنفترض أنَّ كميّات ومتوسط أسعار الطن من هذه المواد والسلع مقدّرة بـ 1000

وحدة نقدية مُعطاة من خلال الجدول الآتي:

| المادة أو السلعة | الكمية | قيمة المادة أو السلعة المصدَّرة | الكمية | قيمة المادة أو السلعة المستوردة |
|---|---|---|---|---|
| القمح | 5000 | 1000 | 6000 | 900 |
| الذرة | 250 | 1250 | 165 | 1100 |
| الملابس | 20 | 5000 | 15 | 6500 |
| الأدوية | 12 | 25000 | 15 | 32000 |

والمطلوب ما يلي:

أ - عيّن دورتي الأساس والمقارنة مع تعليل السبب إن كان هناك تفضيل لأحدهما على الآخر.

ب- احسب المؤشّرات البسيطة للمُعطيات السابقة.

ج- احسب المؤشّر الكلي البسيط للمُعطيات السابقة، وماذا تستنتج من ذلك؟

د- احسب المؤشّر المثالي لـ فيشر من أجل المُعطيات السابقة، وماذا تستنتج من ذلك؟

٤- لنفترض أنَّ إحدى المؤسسات العلمية المهتمّة بالبيئة تقوم بدراسة التلوث في داخل المدن معتمدة في ذلك معايير معيَّنة تحدِّد من خلالها نسبة التلوث، وأنَّ دراستها شملت عدداً من المدن، فإذا قامت بهذه الدراسة في فصلي الشتاء والصيف من عام ما، وأنَّ النتائج كانت كما في الجدول الآتي:

| ر.م. $i$ | الشيء الخاضع للدراسة | في الصيف | | في الشتاء | |
|---|---|---|---|---|---|
| | | عدد المدن الخاضعة للدراسة | نسبة التلوث | عدد المدن الخاضعة للدراسة | نسبة التلوث |
| 1 | الهواء | 9 | 67% | 8 | 75% |
| 2 | الماء | 7 | 12% | 8 | 15% |
| 3 | التربة | 12 | 1.5% | 12 | 2% |
| 4 | الضجيج | 8 | 78% | 9 | 65% |
| Total | ------------- | | | | |

والمطلوب ما يلي:

أ - عيّن دورتي الأساس والمقارنة مع تعليل السبب إن كان هناك تفضيل لأحدهما على الآخر.

ب- احسب المؤشّرات البسيطة للمُعطيات السابقة.

ج- احسب المؤشّر الكلي البسيط للمُعطيات السابقة، وماذا تستنتج من ذلك؟

د- احسب المؤشّر المثالي لـ فيشر من أجل المُعطيات السابقة، وماذا تستنتج من ذلك؟

٥- لنفترض أنَّ أحد المصانع المنتجة للإسمنت تقوم بدراسة حول تأثر الخرسانة المصنَّعة من هذا الاسمنت في مناطق مختلفة بدرجة رطوبتها (تختلف فيها بينها بقيمة متوسط الرطوبة السنوية لها) معتمدة في ذلك معايير معيَّنة لقياس التلف الناتج عن الرطوبة في كل متر مكّعب من الخرسانة بعد مرور فترة من زمن استخدامها (مقدّرة بالسنة)، وأنَّ دراستها شملت عدداً من المناطق، فإذا قامت بهذه الدراسة في فترتين من الزمن (بعد مرور عشر سنوات، و بعد مرور خمسة عشر سنة)، وأنَّ النتائج كانت كما في الجدول الآتي:

| ر.م. $i$ | المنطقة الخاضعة للدراسة | بعد مرور عشر سنوات | | بعد مرور خمسة عشر سنة | |
|---|---|---|---|---|---|
| | | عدد المناطق الخاضعة للدراسة | النسبة المئوية للتلف في المتر المكعب من الخرسانة | عدد المناطق الخاضعة للدراسة | النسبة المئوية للتلف في المتر المكعب من الخرسانة |
| 1 | الساحل | 12 | 23% | 14 | 58% |
| 2 | المرتفعات الجبلية | 10 | 15% | 8 | 37% |
| 3 | الداخلية التي على ارتفاع ما بين 400 وحتى 700 متر عن سطح البحر | 26 | 2% | 32 | 5% |
| Total | ------------- | | | | |

والمطلوب ما يلي:

أ - عيّن دورتي الأساس والمقارنة مع تعليل السبب إن كان هناك تفضيل لأحدهما على الآخر.

ب- احسب المؤشّرات البسيطة للمُعطيات السابقة.

ج- احسب المؤشّر الكلي البسيط للمُعطيات السابقة، وماذا تستنتج من ذلك؟

د- احسب المؤشّر المثالي لـ فيشر من أجل المُعطيات السابقة، وماذا تستنتج من ذلك؟

# الملحق (A)

## رموز ومصطلحات استخدمت في الكتاب

| الرمز | دلالة الرمز ومعناه |
|---|---|
| $:=$ | يعني "يساوي بالتعريف" وقد كان العاملون في مجال الكوبرنيتك هم أول من استخدم هذا الرمز. |
| $\triangleq$ | يعني "يساوي اصطلاحاً"، بمعنى أنَّ ما سيأتي بعده هو تعبير اصطلاحي سنعتمده لما يسبقه من مقادير أو صيغ. |
| $;$ | داخل الأقواس الممثِّلة للمجموعات والحوادث للتعبير عن العبارة "وذلك لكل" أو بمعنى "وذلك من أجل" حسب الدلالة المناسبة لها. |
| $:$ | داخل الأقواس الممثِّلة للمجموعات والحوادث للتعبير عن العبارة "علماً أنَّ". |
| $\mid$ | للتعبير عن الصيغ الشرطية في دراسة الاحتمالات الشـرطية. |
| $\mathbb{N} := \{1,2,3,\ldots\}$ | مجموعة الأعداد الطبيعية Natural Numbers |
| $\mathbb{N}^o := \mathbb{N} \cup \{0\}$ | مجموعة الأعداد الطبيعية مع الصفر Whole Numbers |
| $\bar{\mathbb{N}} := \mathbb{N}^o \cup \{+\infty\}$ | مجموعة الأعداد الطبيعية الموسَّعة Extended Natural Numbers |
| $\mathbb{N}_n := \{1,2,3,\ldots,n\}$ | علماً أنَّ $n \in \mathbb{N}$ كيفي ولكن مُثبَّت. |
| $\mathbb{Z}$ | مجموعة الأعداد الصحيحة Integer Number |
| $\mathbb{Q}$ | مجموعة الأعداد النسبيّة Rational Numbers |
| $\mathbb{R}$ | مجموعة الأعداد الحقيقية Real Number |
| $\mathbb{R}^* := \mathbb{R} \setminus \{0\}$ | مجموعة الأعداد الحقيقية ما عدا الصفر |
| $\mathbb{R}^+ := \mathbb{R} \setminus (-\infty,0]$ | مجموعة الأعداد الحقيقية الموجبة |

| الرمز | دلالة الرمز ومعناه |
| --- | --- |
| $\bar{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ | مجموعة الأعداد الحقيقية الموسَّعة |
| $\mathbb{C}$ | مجموعة الأعداد العقدية (أو المركَّبة) Complex Number |
| $\lceil x \rceil$ | يشير إلى أصغر عدد صحيح أكبر أو يساوي $x$، فعلى سبيل المثال: $\lceil 8.99 \rceil = 9$ & $\lceil 8.11 \rceil = 9$ & $\lceil 8 \rceil = 8$ |
| $\lfloor x \rfloor$ | يشير إلى أكبر عدد صحيح أصغر أو يساوي $x$، فعلى سبيل المثال: $\lfloor 8.99 \rfloor = 8$ & $\lfloor 8.11 \rfloor = 8$ $\lfloor 8 \rfloor = 8$ |
| $a << b$ | يعني أنَّ $b$ أكبر بكثير من $a$، علماً أنَّ $a$ و $b$ عددين حقيقيين. |
| $1 << a$ | يعني أنَّ $a$ كبيراً بقدرٍ كافٍ، علماً أنَّ $a$ عدد حقيقي موجب. |
| $1 >> a$ | يعني أنَّ $a$ صغيراً بقدرٍ كافٍ، علماً أنَّ $a$ عدد حقيقي موجب. |
| $\wedge$ | الرمز المنطقي الذي يشير إلى عبارة ( و and). |
| $\vee$ | الرمز المنطقي الذي يشير إلى عبارة ( أو or). |
| $\circ$ | يشير إلى تركيب التطبيقات (أو الدوال). |
| $\lambda \prec\prec \mu$ | يعني الاستمرار المطلق بين قياسين $\lambda$ و $\mu$ فوق فضاء مقيس. |
| $\mathfrak{Z}_2 )) \mathfrak{Z}_1$ | يعني أنَّ التجزئة $\mathfrak{Z}_1$ أنعم من التجزئة $\mathfrak{Z}_2$ . |

## مفاهيم عامة من الرياضيات

نقدّم فيما يلي بعض المفاهيم والمصطلحات الرياضياتية التي استخدمت في فصول هذا الكتاب وذلك بغية التذكير بها، لأنّه من المفترض أن يكون القارئ ملمًّا بها قبل دراسة هذا الكتاب. إنَّ البراهين والإثباتات الخاصّة بفقرات هذا الملحق تمَّ تجاوزها حيث يمكن لمن يودّ الاطلاع عليها الرجوع إلى المراجع ذات الصلة ومنها [1] و [16] و [23] في فهرس المراجع لهذا الكتاب.

### A-١ - المجموعات

من المعلوم أنَّ مفهوم المجموعة يُعدّ من الأهمية بمكان في مجال الرياضيات عامّة بحيث إنَّه يأتي بالمرتبة الثانية من حيث الأهمية بعد مفهوم الأعداد، ولكنَّنا سوف نستعرض (وبشكل موجز) جانباً يسيراً من نظرية المجموعات مع التركيز على المواضيع التي كانت موضع اهتمامنا في هذا الكتاب، علماً أنَّنا سنستخدم ما اتفق عليه فقط دون الخوض في تفصيلات المسائل وبراهينها.

#### A-١-١ - تعريف (المجموعة Set)

المجموعة هي أي تجمّع لأشياء مُحدَّدة ومتمايزة جيداً بحسب حدسنا أو تفكيرنا، وصُبَّت في قالب موحَّد.

هذا التعريف للمجموعة كان قد قُدِّم من قبل الرياضياتي الألماني كانتور (1845-1918) Georg Cantor الذي أسس نظرية المجموعات Sets Theory.

إنَّ الأشياء المكوِّنة للمجموعة تُدعى عناصر، وقد درجت العادة على أن يُرمز للمجموعات بحروف لاتينية كبيرة من قبيل $A$، $B$، $C$ و...، أو بأحرف إغريقية كبيرة من قبيل $\Omega$، $\Lambda$، $\Theta$ و...، وأمّا عناصر المجموعات فيرمز لها بأحرف لاتينية صغيرة من قبيل $a$، $b$، $c$ و...، أو بأحرف إغريقية صغيرة من قبيل $\alpha$، $\beta$، $\gamma$ و...

في الحقيقة إنَّ التعريف السابق هو التعريف العام للمجموعة، حيث يُطلق على المجموعة المنوّه عنها آنفاً اسم المجموعة الشاملة، ولكن توجد طرائق عديدة لعرض المجموعات يُفترض أن تكون معلومة لدى القارئ (لأنَّها تدرّس في مرحلة التعليم قبل الجامعي)، ولذلك سنفترض أنَّ القارئ ملمٌّ بالقضايا المتعلّقة بالمجموعات (مثل طرائق عرضها، والعمليّات الجبرية على المجموعات، وطرائق تمثيلها بيانياً و ...) التي تدرَّس في مراحل التعليم السابقة للتعليم الجامعي (مرحلة التعليم الأساسي والثانوية)، وما سنقوم بتقديمه هنا سيكون على سبيل التذكير فقط.

الآن بفرض أنَّ $\Omega$ مجموعة شاملة، فعندئذ مجموعة العناصر $A$ من $\Omega$ التي تتمتَّع بصفة مُحدَّدة تدعى مجموعات جزئية، وفي حال كانت هذه المجموعة لا تحوي أي عنصر فحينئذ تُدعى مجموعة خالية ويرمز لها بـ $\varnothing$، ومن ثمَّ نلاحظ أنَّ المجموعة الخالية محتواة في أي مجموعة. كذلك بفرض أنَّ $B \subseteq \Omega$ عندئذ يطلق على مجموعة العناصر التي تنتمي إلى $\Omega$ ولا تنتمي إلى $B$ اسم المجموعة المتمِّمة لـ $B$ ويُرمز لها بـ $\overline{B}$.

#### A-١-٢ - تعريف (قدرة مجموعة Cardinality of a Set)

لتكن $A$ مجموعة ما من مجموعة شاملة $\Omega$، فعندئذ تُعَرَّف قدرة $A$ على أنَّها عدد عناصرها، ويرمز لهذا العدد بـ $|A|$، فإذا كان لدينا على سبيل المثال $A = \{a_1, a_2, ..., a_n\}$ مع $n \in \mathbb{N}$ مثَّبت، فإنَّه سيكون لدينا $|A| = n$، وأما إذا كانت $A$ غير منتهية فعندئذ نميِّز بين حالتين:

**الأولى**: إذا كانت $A$ غير منتهية ولكن يمكن إيجاد تطبيق تقابل $\sigma$ من $\mathbb{N}$ على $A$، فعندئذ يُقال إنَّ المجموعة $A$ **قابلة للعدّ** Countable Set، ويُكتب $|A| = \infty$. إنَّ التعبير السابق يعني أنَّه يمكن ترقيم عناصر المجموعة $A$ بوساطة الأعداد الطبيعية $\mathbb{N}$.

**الثانية**: إذا كانت $A$ غير منتهية ولكن يمكن إيجاد تطبيق تقابل $\sigma$ من $\mathbb{R}$ على $A$، فعندئذ يُقال إنَّ المجموعة $A$ **غير قابلة للعدّ** Uncountable Set، ويُطلق على قدرتها اسم **قدرة المستمر** ورمزها $\wp$ أي إنَّ $|\mathbb{R}| = \wp$، ولذلك يُكتب $|A| = \wp$.

بناء على ما سبق يمكننا أن نذكّر بالآتي:

- إذا وجد تطبيق تقابل $\sigma$ بين مجموعتين $A$ و $B$، فعندئذ يُقال إنَّ $A$ **تكافئ** $B$ أو إنَّ $A$ و$B$ **متكافئتان**، ويُكتب حينئذ $A \sim B$.
- إذا حُذفت مجموعة منتهية من مجموعة قابلة للعدّ فإنَّ الباقي هي مجموعة قابلة للعدّ أيضاً.
- إذا حُذفت مجموعة منتهية من مجموعة غير قابلة للعدّ فإنَّ الباقي هي مجموعة غير قابلة للعدّ أيضاً.
- إذا حُذفت مجموعة قابلة للعدّ من مجموعة غير قابلة للعدّ فإنَّ الباقي هو مجموعة غير قابلة للعدّ أيضاً.
- إنَّ المجموعة الناتجة عن اتحاد (أو اجتماع) عدد منتهٍ لمجموعات قابلة للعدّ هي من جديد مجموعة قابلة للعدّ.
- لتكن $A$ مجموعة غير منتهية ولكنَّها قابلة للعدّ على الأكثر، فعندئذ تكون أسرة كل المجموعات الجزئية من $A$ (ويُرمز لها بـ $2^A$) هي مجموعة غير قابلة للعدّ.
- إذا كانت $A_1$، $A_2$، ...و $A_n$ مجموعات قابلة للعدّ مع $n \in \mathbb{N}$ كيفي ولكن مثبّت، فإنَّ المجموعة $B = A_1 \times A_2 \times \cdots \times A_n$ تكون قابلة للعدّ أيضاً.

### A-١-٣ - العمليّات الجبرية على المجموعات

لتكن $A$ و$B$ مجموعتين من مجموعة شاملة $\Omega$، فعندئذ:

١ - **اتحاد** المجموعتين $A$ و$B$ هي مجموعة كل العناصر التي تنتمي إلى $A$ أو إلى $B$ أو إلى كليهما معاً، ويرمز لها ب $A \cup B$، علماً أنَّ عبارة (أو) هنا لا تفيد الحصر. أي إنَّه لدينا:

$$A \cup B = \{\omega \in \Omega \,;\, \omega \in A \quad or \quad \omega \in B\}$$

وإذا كانت المجموعات $A_1$، $A_2$، ...و $A_n$ منفصلة مثنى مثنى فإنَّه يمكن استخدام الرمز $\sum_{i=1}^{n} A_i$ بدلاً من $\bigcup_{i=1}^{n} A_i$ للدلالة على أنَّ المجموعات التي تحت رمز الاتحاد منفصلة مثنى مثنى.

٢- **تقاطع** المجموعتين $A$ و$B$ هي مجموعة كل العناصر التي تنتمي إلى كل من $A$ و$B$ معاً، ويرمز لها بـ $A \cap B$. أي إنَّه لدينا:

$$A \cap B = \{\omega \in \Omega \,;\, \omega \in A \quad and \quad \omega \in B\}$$

٣- **الفرق** النّاتج عن طرح المجموعة $A$ من $B$ (ويُقال $A$ ناقص $B$) هي مجموعة كل العناصر التي تنتمي إلى $A$ ولا تنتمي إلى $B$، ويُرمز لها $A \setminus B$. أي إنَّه لدينا:

$$A \setminus B := \{\omega \in \Omega \,;\, \omega \in A \quad and \quad \omega \notin B\}$$

٤- **الفرق التناظري** بين $A$ و$B$ هي مجموعة كل العناصر من $A$ و$B$ ولا تنتمي إلى $A \cap B$، ويُرمز له بـ $A \,\Delta\, B$، أي إنَّه لدينا:

$$A \Delta B = (A \cup B) \setminus (A \cap B)$$

الآن، لتكن $\Omega$ مجموعة شاملة، ولنفترض أنَّ $A$، $B$، و$C$ مجموعات جزئية من $\Omega$، فعندئذ يمكن للقارئ التّحقُّق من صحة

العلاقات الآتية:

**a)** $A \cup B = B \cup A \quad \& \quad A \cap B = B \cap A$

**b)** $(A \cup B) \cup C = A \cup (B \cup C) \quad \& \quad (A \cap B) \cap C = A \cap (B \cap C)$

**c)** $A \cup (B \cap C) = (A \cup B) \cap (A \cup C) \quad \& \quad A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

**d)** $A \cap (B \setminus C) = (A \cap B) \setminus (A \cap C) = (A \cap B) \setminus C$

**e)** $\overline{A \cup B} = \bar{A} \cap \bar{B} \quad \& \quad \overline{A \cap B} = \bar{A} \cup \bar{B}$ قانوني دي مورغان

وهاتين العلاقتين الأخيرتين تبقيان ساريتا المفعول من أجل أي عدد منته أو غير منته (ولكن قابل للعدّ) من المجموعات، أي إنَّه من أجل أي متتالية مجموعات $A_1$، $A_2$، ...و $A_n$ و... من مجموعة شاملة $\Omega$، يكون لدينا:

**f)** $\overline{\bigcup_{n=1}^{\infty} A_n} = \bigcap_{n=1}^{\infty} \overline{A_n} \quad \& \quad \overline{\bigcap_{n=1}^{\infty} A_n} = \bigcup_{n=1}^{\infty} \overline{A_n}$

**g)** $A \setminus B = A \cap \bar{B} \quad \& \quad A \setminus \bar{B} = A \cap B \quad \& \quad (A \setminus B) \cup (B \setminus A) = (A \cap \bar{B}) \cup (B \cap \bar{A})$

وكذلك إذا كانت $A \supseteq B$ فعندئذ يكون لدينا:

$$\textbf{h)} \quad \begin{cases} \bar{B} \subseteq \bar{A} & \& \quad A \cap B = A \\ A \cup B = B & \& \quad B \setminus A = \bar{A} \setminus \bar{B} \end{cases}$$

### A-١-٤ - متتاليات المجموعات المطّردة:

لتكن $(A_n)_{n\in\mathbb{N}}$ متتالية مجموعات من مجموعة شاملة $\Omega$ ، فإذا كان:

١- لدينا $A_1 \supseteq A_2 \supseteq ... \supseteq A_n \supseteq ...$ فإنَّه يُقال عن المتتالية $(A_n)_{n\in\mathbb{N}}$ إنَّها متتالية مجموعات **غير متناقصة**.

٢- لدينا $A_1 \supset A_2 \supset ... \supset A_n \supset ...$ فإنَّه يُقال عن المتتالية $(A_n)_{n\in\mathbb{N}}$ إنَّها متتالية مجموعات **متزايدة**.

٣- لدينا $A_1 \subseteq A_2 \subseteq ... \subseteq A_n \subseteq ...$ فإنَّه يُقال عن المتتالية $(A_n)_{n\in\mathbb{N}}$ إنَّها متتالية مجموعات **غير متزايدة**.

٤- لدينا $A_1 \subset A_2 \subset ... \subset A_n \subset ...$ فإنَّه يُقال عن المتتالية $(A_n)_{n\in\mathbb{N}}$ إنَّها متتالية مجموعات **متناقصة**.

علماً أنَّه إذا وافقت المتتالية $(A_n)_{n\in\mathbb{N}}$ إحدى الحالات الأربع السابقة فعندئذ يُقال عن هذه المتتالية إنَّها **مطَّردة**.

### A-١-٥ - تقارب متتاليات المجموعات

لتكن $(A_n)_{n\in\mathbb{N}}$ متتالية مجموعات من مجموعة شاملة $\Omega$ ، فعندئذ:

١- إذا كانت $(A_n)_{n\in\mathbb{N}}$ غير متناقصة (**أو متزايدة تماماً**) فإنَّ نهاية هذه المتتالية (**ولتكن مجموعة** $A$) هو اتحاد عناصرها، أي إنَّ:

$$A = \lim_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} A_n \quad \textbf{or} \quad A = \lim_{n\uparrow\infty} A_n = \bigcup_{n=1}^{\infty} A_n$$

٢- إذا كانت $(A_n)_{n\in\mathbb{N}}$ غير متزايدة (**أو متناقصة تماماً**) فإنَّ نهاية هذه المتتالية (**ولتكن مجموعة** $A$) هو تقاطع عناصرها، أي إنَّ:

$$A = \lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n \quad \textbf{or} \quad A = \lim_{n\downarrow\infty} A_n = \bigcap_{n=1}^{\infty} A_n$$

٣- تُعرَّف النهاية العليا للمتتالية $(A_n)_{n\in\mathbb{N}}$ (في حال وجودها) على أنَّها المجموعة $A=\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k$، ويكتب حينئذ:

$$A=\overline{\lim_{n\to\infty}}\,A_n=\limsup A_n=\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty}A_k$$

٤- تُعرَّف النهاية الدنيا لهذه المتتالية $(A_n)_{n\in\mathbb{N}}$ (في حال وجودها) على أنَّها المجموعة $A=\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}A_k$، ويُكتب حينئذ:

$$A=\underline{\lim_{n\to\infty}}\,A_n=\liminf A_n=\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}A_k$$

وبناءً على ما سبق يمكن البرهنة على صحة العلاقات الآتية:

a) $\overline{\limsup A_n}=\liminf\overline{A_n}$ & $\overline{\liminf A_n}=\limsup\overline{A_n}$

b) $\liminf A_n\subseteq\limsup A_n$

c) $\limsup\,(A_n\cup B_n)=\limsup A_n\cup\limsup B_n$

d) $\limsup A_n\cap\liminf B_n\subseteq\limsup(A_n\cap B_n)\subseteq\limsup A_n\cap\limsup B_n$

علماً أنَّ الخط العلوي يَرمز إلى المُتمِّم، وفي الحالة الخاصّة عندما تكون المتتالية $(A_n)_{n\in\mathbb{N}}$ مطَّردة ومتقاربة من مجموعة $A$ فإنَّه يكون لدينا:

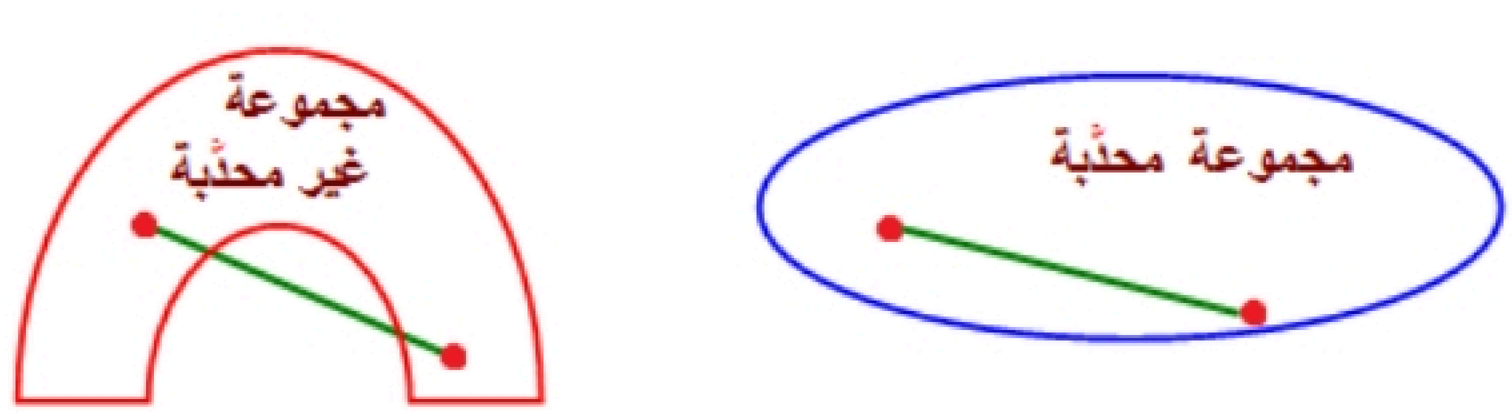
e) $\liminf A_n=\limsup A_n=A$

### A-١-٥ - تعريف (المجموعة المحدَّبة Convex Set)

يُقال عن مجموعة $D\subseteq\mathbb{R}^n$ إنَّها مُحدَّبة إذا كان من أجل كل $\vec{x}_n=(x_1,x_2,\dots,x_n)$ و $\vec{y}_n=(y_1,y_2,\dots,y_n)$ من $D$، ومن أجل أي $0\le\theta\le 1$ لدينا العلاقة الآتية محقَّقةً:

$$\theta\cdot\vec{x}_n+(1-\theta)\cdot\vec{y}_n\in D$$

وفي الحالة الخاصّة لو أخذنا $D\subseteq\mathbb{R}^2$ فإنَّه إذا كان المستقيم الواصل بين أيّة نقطتين من $D$ يقع بكامله في $D$ أيضاً، فعندئذ يُقال عن المجموعة $D$ إنَّها مُحدَّبة.




## A-٢ - مفاهيم من التحليل المزجي (التوافقي)

### A-٢-١ - تعريف (التباديل Permutations)

لتكن $A$ و $B$ مجموعتين منتهيتين قدرة كل منهما تساوي $n$، فعندئذ كل تطبيق تقابل $\pi: A\longleftrightarrow B$ يُدعى تبديلاً لعناصر المجموعة $B$ وفقاً لعناصر المجموعة $A$.

إنَّ مجموعة كل تباديل عناصر $B$ وفقاً لعناصر $A$ سنرمز لها بـ $\Pi_{A,B}$، ومن أجل $A=B$ سنكتب $\Pi_A$ عوضاً عن $\Pi_{A,A}$، وفي الحالة الخاصّة عندما يكون لدينا $A=\mathbb{N}_n$ مع $n\in\mathbb{N}$ فإنَّنا سنكتب $\Pi_n$ عوضاً عن $\Pi_{\mathbb{N}_n}$.

بناءً على التعريف السّابق يكون من أجل $|A|=|B|=n$ لدينا $|\Pi_{A,B}|=n\,!$، ومن المفيد في كثير من الحالات استخدام **نشر سترلنغ** Stirling Expansion (نسبة إلى الرياضي الأسكتلندي سترلنغ **James Stirling (1692-1770)**) لحساب المقدار $n\,!$ باستخدام العلاقة الآتية:

$$n\,!=\sqrt{2n\pi}\;n^n\,\mathrm{e}^{-n}\,\mathrm{e}^{\frac{\theta}{12n}}\qquad ;n\in\mathbb{N}\;\;\&\;\;0<\theta<1$$

وعندما تكون $n$ كبيرة بقدر كافٍ يُمكننا أن نكتب وبتقريب جيد ما يلي:

$$n\,!\approx\sqrt{2\pi n}\;\;n^n\;\;\mathrm{e}^{-n}\qquad ;n\in\mathbb{N}$$

إذا وجدت في المجموعة $B$ عناصر متماثلة كأن يكون فيها $n_1$ عنصر من نوع أول، و $n_2$ عنصر من نوع ثاني، و... و $n_k$ عنصر من نوع أخير، وبحيث يكون $\sum_{i=1}^{k}n_i=n$، فعندئذ يكون عدد التباديل في المجموعة $\Pi_{A,B}$ يساوي:

$$\binom{n}{n_1,n_2,\cdots\cdots,n_k}\;\triangleq\;\frac{n\,!}{n_1\,!\cdot n_2\,!\cdot\ldots\ldots\cdot n_k\,!}$$

يُقال عن تبديل $\boldsymbol{\pi}\in\Pi_{\mathbb{N}}$ إنَّه **تبديل منته** Finite Permutation إذا وجد عدد منته من القيم $j\in\mathbb{N}$ من أجلها لدينا $\boldsymbol{\pi}(j)\neq j$. أي إنَّه من أجل كل $n\in\mathbb{N}$ لدينا $\boldsymbol{\pi}(n)=n$ باستثناء عدد منته من هذه القيم $n$.

### A-٢-٢ - تعريف (التراتيب)

لتكن $A$ و $B$ مجموعتين منتهيتين بحيث إن $|A|=k$ و $|B|=n$ مع $n\geq k$، فعندئذ يُسمَّى كل تطبيق متباين (تطبيق واحد لواحد) $\zeta:A\to B$ ترتيباً لعناصر المجموعة $B$ وفقاً لعناصر المجموعة $A$، فإذا رمزنا لمجموعة تراتيب المجموعة $B$ وفقاً لعناصر $A$ بـ $\vartheta_{A,B}$، فعندئذ يكون لدينا:

$$\left|\vartheta_{A,B}\right|=n\cdot(n-1)\cdot\ldots\ldots\cdot[n-(k-1)]=\frac{n\,!}{(n-k)!}\triangleq n\,\mathbf{P}\,k$$

وفي الحالة الخاصّة عندما $k=n$ فإنَّ التراتيب تؤول إلى التباديل التي قُدِّمت سابقاً ويكون لدينا في هذه الحالة $n\,\mathbf{P}\,n=n\,!$.

### A-٢-٣ - تعريف (التوافيق Combinations)

لتكن $A$ و $B$ مجموعتين منتهيتين بحيث $|A|=k$ و $|B|=n$ مع $n\geq k$، وليكن $\beta:A\to B$ تطبيقاً متبايناً، فإذا كان المتَّجه $\mathbb{V}\in B\times B\times\cdots\times B=B^k$ يُحقِّق العلاقة الآتية من أجل أيّ تبديل $\boldsymbol{\pi}\in\Pi_k$:

$$\left(\beta(a_{\boldsymbol{\pi}(1)})\,,\,\beta(a_{\boldsymbol{\pi}(2)})\,,\ldots,\,\beta(a_{\boldsymbol{\pi}(k)})\right)=\left(\beta(a_1)\,,\,\beta(a_2)\,,\ldots,\,\beta(a_k)\right)=\mathbb{V}$$

فعندئذ يُسمَّى التطبيق $\beta$ توفيقاً لعناصر المجموعة $B$ وفقاً لعناصر المجموعة $A$.

سنرمز لمجموعة كل التوافيق لعناصر $B$ وفقاً لعناصر $A$ بـ $\mathcal{C}_{A,B}$، وإذا كان لدينا $|A|=k$ و $|B|=n$ مع $n\geq k$، فإنَّه سيكون لدينا:

$$\left|\mathcal{C}_{A,B}\right|=\frac{n\,!}{k\,!\,(n-k)!}\;\triangleq\;\binom{n}{k}$$

وذلك بسبب إزالة خاصيّة أهميّة الترتيب التي عدد حالاتها الممكنة يساوي $k\,!$، وفي الحالة الخاصّة عندما $k=n$ فإنَّ مفهوم التوافيق يفقد معناه، وذلك لأنَّه سيصبح لدينا $\left|\mathcal{C}_{A,B}\right|=1$.

## A-٣ - القاعدتان الأساسيتان للعدّ:

### A-٣-١ - قاعدة الضرب Multiplication Rule:

لتكن $A$ و $B$ مجموعتين منتهيتين بقدرة $n$ و $m$ على الترتيب، فعندئذ تكون قدرة المجموعة الناتجة عن الضرب (أو الجداء) الديكارتي لهما $A\times B$ تساوي $n\cdot m$، وهذه النتيجة يمكن تعميمها على أي عدد منته من المجموعات المنتهية، فإذا كان لدينا $A_1$، $A_2$، ...، و $A_k$ مجموعات منتهية بقدرات $n_1$، $n_2$، ...، و $n_k$ على الترتيب، فإنَّ قدرة المجموعة $A_1\times A_2\times\ldots\times A_k$ ستكون مساوية إلى $n_1\cdot n_2\cdot\ldots\cdot n_k$، أي إنَّه لدينا $\left|\underset{i=1}{\overset{k}{\times}} A_i\right|=\prod_{i=1}^{k} n_i$، و هذه القاعدة تُعرف باسم **قاعدة الضرب**.

### A-٣-١ - قاعدة الجمع Addition Rule

لتكن $A$ و $B$ مجموعتين منتهيتين بقدرات $n$ و $m$ على الترتيب، فإذا كان لدينا $A\cap B=\varnothing$ فإنَّ قدرة المجموعة $A\cup B$ ستكون مساوية لـ $n+m$، وهذه النتيجة يمكن تعميمها على أي عدد منته من المجموعات المنتهية، فإذا كانت $A_1$، $A_2$، ...، و $A_k$ مجموعات غير خالية ومنفصلة مثنى مثنى بقدرات $n_1$، $n_2$، ...، و $n_k$ على الترتيب، فإنَّ قدرة المجموعة $B=\sum_{i=1}^{k} A_i$ تساوي $\sum_{i=1}^{k} n_i$، أي إنَّه لدينا $\left|\sum_{i=1}^{k} A_i\right|=\sum_{i=1}^{k} n_i$، إنَّ هذه القاعدة تُعرف باسم **قاعدة الجمع**.

## A-٤ دوال بمواصفات خاصَّة

أخيراً نقدِّم بعض الدوال الخاصّة التي ورد ذكرها في دراستنا لهذا الكتاب.

### A-٤-١ - تعريف ( الدالَّة المميِّزة لمجموعة (Indicator Function of a Set

لتكن $\Omega$ مجموعةً ما غير خالية، ولنأخذ $A\subseteq\Omega$ مجموعة جزئية، ولنعرِّف من أجل هذه المجموعة دالَّة حقيقية $\mathbf{I}_A$ على $\Omega$ كما يلي:

$$\mathbf{I}_A(\omega)=\begin{cases}1 & for \quad \omega\in A\\ \\ 0 & for \quad \omega\notin A\end{cases}$$

فعندئذ تُدعى $\mathbf{I}_A$ **الدالَّة المميِّزة للمجموعة $A$**، وذلك لأنَّها تميِّز عناصر المجموعة $A$ بشكل تام.

### A-٤-١-١ - خصائص الدالَّة المميِّزة لمجموعة

نظراً لأهمّية الدالَّة $\mathbf{I}_A$ سنقدِّم بعض خصائصها حيث يمكن للقارئ التَحقُّق من صحتها:

**a)** $\mathbf{I}_\Omega(\omega)=1 \quad ;\forall\omega\in\Omega$

**b)** $\mathbf{I}_\varnothing(\omega)=0 \quad ;\forall\omega\in\Omega$

**c)** For $A, B\subseteq\Omega$ we have $\mathbf{I}_{A\cap B}(\omega)=\mathbf{I}_A(\omega)\cdot\mathbf{I}_B(\omega) \quad ;\forall\omega\in\Omega$

**d)** For $A, B\subseteq\Omega$ we have $\mathbf{I}_{A\cup B}(\omega)=\mathbf{I}_A(\omega)+\mathbf{I}_B(\omega)-\mathbf{I}_{A\cap B}(\omega) \quad ;\forall\omega\in\Omega$

**e)** For $A, B\subseteq\Omega$ we have $\mathbf{I}_{A\Delta B}(\omega)=\left(\mathbf{I}_A(\omega)-\mathbf{I}_B(\omega)\right)^2 \quad ;\forall\omega\in\Omega$

### A-٤-٢- تعريف (الدالَّة المحدَّبة Convex Function)

لتكن $f: D \to \mathbb{R}$ دالَّة مُعطاة مع $D \subseteq \mathbb{R}^n$ مجموعة محدَّبة، فإذا كان من أجل أي $0 \le \theta \le 1$ وكل $\vec{x}_n = (x_1, x_2, \ldots, x_n)$ و $\vec{y}_n = (y_1, y_2, \ldots, y_n)$ من $D$ لدينا العلاقة الآتية مُحقَّقة:

$$f\left(\theta \cdot \vec{x}_n + (1-\theta) \cdot \vec{y}_n\right) \le \theta \cdot f(\vec{x}_n) + (1-\theta) \cdot f(\vec{y}_n)$$

فعندئذ يُقال عن الدالَّة $f$ إنَّها **محدَّبة**، وإذا كانت العلاقة الآتية محقَّقة من أجل كل $\vec{x}_n \ne \vec{y}_n$:

$$f\left(\theta \cdot \vec{x}_n + (1-\theta) \cdot \vec{y}_n\right) < \theta \cdot f(\vec{x}_n) + (1-\theta) \cdot f(\vec{y}_n)$$

فعندئذ يُقال عن الدالَّة $f$ إنَّها محدَّبة على نحو صارم.

من النماذج الشهيرة للدوال الحقيقية المعرَّفة على $D \subseteq \mathbb{R}$ والتي تتمتَّع بخاصَّة التحدُّب نورد الدوال الآتية:

- الدوال الخطية $f(x) = a + bx$ مع $a$ و $b$ من $\mathbb{R}$.
- الدوال الأسيّة $f(x) = \mathbf{e}^{bx}$ مع $b$ من $\mathbb{R}$.
- دوال القوى $f(x) = x^{\alpha}$ من أجل $D \subseteq \mathbb{R}^+$ و $1 \le \alpha$ أو $\alpha \le 0$.
- دوال القيمة المطلقة $f(x) = |x|^p$ من أجل $1 \le p$.

### A-٤-٣ - تعريف (الدالَّة المقعَّرة Concave Function)

يقال عن دالَّة حقيقية $f$ معرَّفة على مجموعة $D \subseteq \mathbb{R}^n$ إنَّها **مقعَّرة** إذا كانت الدالَّة $-f$ محدَّبة، وفي حال كانت الدالَّة $-f$ محدَّبة على نحو صارم، فعندئذ يُقال عن الدالَّة $f$ إنَّها مقعَّرة على نحو صارم.

من النماذج الشهيرة للدوال الحقيقية المعرَّفة على $D \subseteq \mathbb{R}$ التي تتمتَّع بخاصَّة التقعُّر نورد الدوال الآتية:

- الدوال الخطية $f(x) = a + bx$ مع $a$ و $b$ من $\mathbb{R}$، ومن ثمَّ نلاحظ أنَّ الدوال الخطّيّة هي دوال محدَّبة ومقعَّرة بآن واحد.
- دوال القوى $f(x) = x^{\alpha}$ من أجل $D \subseteq \mathbb{R}^+$ و $0 \le \alpha \le 1$.
- الدوال اللوغاريتمية $f(x) = \mathbf{log}\, x$ من أجل $D \subseteq \mathbb{R}^+$.

### A-٤-٤ - تعريف ( الدالَّة الدرجيّة Step Function)

لتكن $f: D \to \mathbb{R}$ دالَّة مُعطاة مع $D \subseteq \mathbb{R}$، فإذا كان من أجل $n \in \mathbb{N}$ (**مثبَّت**) يوجد $a_1$، $a_2$،...، و $a_n \in \mathbb{R}$ و $A_1$، $A_2$، ...، و $A_n \in D$ بحيث يكون للدالَّة $f$ العرض الآتي:

$$f(x) = \sum_{k=1}^{n} a_k \cdot \mathbf{I}_{A_k}(x)$$

فعندئذ يُقال عن الدالَّة $f$ إنَّها **دالَّة درجيّة** Step Function أو **دالَّة بسيطة** Simple Function.

### A-٤-٥ - بعض الخصائص الجبرية للدوال الدرجيّة

١- بفرض أنَّ $f$ و $g$ دالَّتان درجيتان معرَّفتان على مجموعة $D \subseteq \mathbb{R}$، فعندئذ:

أ- من أجل أي $\alpha \in \mathbb{R}^*$ تكون الدالَّة $\alpha \cdot f$ درجيّة على $D$ أيضاً.

ب- تكون الدوال $f \pm g$ و $f \cdot g$ هي دوال درجيّة على $D$ أيضاً.

٢- من أجل أيّة دالّة حقيقية $f$ معرَّفة ومطّردة على المجموعة $D$ توجد متتالية دوال درجيّة $\left(f_n\right)_{n\in\mathbb{N}}$ معرَّفة على $D$ بحيث يكون لدينا:

$$f(x) = \lim_{n\to\infty} f_n(x) \quad ;\forall x \in D$$

٣- من أجل أيّة دالّة حقيقية $f$ معرَّفة ومستمرَّة على المجموعة $D$ توجد متتالية دوال درجيّة $\left(f_n\right)_{n\in\mathbb{N}}$ معرَّفة على $D$ بحيث يكون لدينا:

$$f(x) = \lim_{n\to\infty} f_n(x) \quad ;\forall x \in D$$

**A-٤-٦ - تعريف (دالة غاما Gama Function)**

لتكن $\Gamma$ دالّة حقيقية معرَّفة على الفترة $(0\ ,\ \infty)$ من خلال:

$$\Gamma(x) := \int_0^{+\infty} y^{x-1}\,\mathbf{e}^{-y}\ dy$$

فعندئذ هذه الدالّة تُدعى **دالّة غاما**.

يمكن البرهنة على أنَّ هذه الدالّة تتمتع بالخصائص الآتية:

- الدالّة $\Gamma(x)$ مستمرة وقابلة للاشتقاق عدداً كيفياً من المرّات على مجالها (مجموعة تعريفها)، ومن أجل كل $x\in(0\ ,\ \infty)$ لدينا العلاقة الآتية محقّقةً:

$$\Gamma^{(n)}(x) := \int_0^{+\infty} y^{x-1}\,\mathbf{ln}^n y\ \mathbf{e}^{-y}\ dy$$

- من أجل كل $x\in(0\ ,\ \infty)$ وكل $n\in\mathbb{N}$ لدينا:

$$\Gamma(x+n) = (x+n-1)\cdot(x+n-2)\cdot\ldots\cdot(x+1)\cdot x\cdot\Gamma(x)$$

والتي ينتج عنها أنَّ $\Gamma(n+1) = n!$، ومن ثمَّ يكون لدينا $0! = \Gamma(1) = 1$ وكذلك $1! = \Gamma(2) = 1$.

- من أجل كل $x\in(0\ ,\ \infty)$ لدينا:

$$\Gamma(x)\cdot\Gamma(1-x) = \frac{\pi}{\sin \pi x}$$

والتي ينتج عنها أنَّ $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

- من أجل كل $x\in(0\ ,\ \infty)$ لدينا:

$$\Gamma(x) = 2\int_0^{+\infty} t^{2x-1}\ \mathbf{e}^{-t^2}\ dt$$

ومن النتائج المهمّة للعلاقة السابقة النتيجة:

$$\int_0^{+\infty} \mathbf{e}^{-t^2}\,dt = \frac{\Gamma(1/2)}{2} = \frac{\sqrt{\pi}}{2}$$

### A-٤-٧ - تعريف (دالة بيتا Beta Function)

لتكن $\beta$ دالّة حقيقية معرَّفة على المجال $(0 , \infty)\times(0 , \infty)$ كما يلي:

$$\beta (p , q) := \int_0^1 y^{p-1} (1-y)^{1-q} \, dy$$

فعندئذ هذه الدالّة تُدعى **دالّة بيتا**.

يمكن البرهنة على أنّه من أجل كل $(p , q)$ من المجال $(0 , \infty)\times(0 , \infty)$ فإنَّ دالّة بيتا تتمتع بالخصائص الآتية:

١- إذا كان $p > 1$ فإنَّه سيكون لدينا:

$$\beta (p , q) = \frac{q-1}{p+q-1} \beta (p , q-1)$$

٢- إذا كان $q > 1$ فإنَّه سيكون لدينا:

$$\beta (p , q) = \frac{p-1}{p+q-1} \beta (p-1 , q)$$

٣- من أجل كل $n, m \in \mathbb{N}$ يكون لدينا:

$$\beta (n , m) = \frac{(n-1)! \, (m-1)!}{(n+m-1)!}$$

٤- من أجل كل $(p , q) \in (0 , \infty)\times(0 , \infty)$ يكون لدينا:

$$\beta (p , q) = \frac{\Gamma(p) \cdot \Gamma(p)}{\Gamma(p+q)}$$

## بهذا نختم الملحق $A$

# الملحق (B)

## توليد جبر الحوادث وبعض المفاهيم من نظرية القياس

سنقدّم في هذا الملحق تذكيراً بموضوعات تتعلّق بنظرية القياس آخذين في الحسبان الرموز، والمصطلحات، والمفاهيم التي قدّمت في هذاالكتاب، وكذلك نشير إلى أنّنا لن نخوض في تفصيلات هذه الموضوعات ولن نقدّم أمثلة عليها إلاّ إذا اقتضت الضرورة ذلك حيث يمكن لمن يودّ الاطلاع على أمثلة وتوضيحات إضافية الرجوع إلى المراجع ذات الصلة ومنها [4] و[5] و[8] و[31] و[39] و[42] و[53] و[56] في فهرس المراجع لهذا الكتاب.

من أجل مجموعة $\Omega$ ما غير خالية لدينا $2^{\Omega}$ أسرة (أو صف) كل المجموعات الجزئية من مجموعة $\Omega$، ومن أجل $A \subseteq \Omega$ تكون المجموعة $\overline{A}$ هي متمّمة المجموعة $A$ بالنسبة إلى $\Omega$.

### B-١ - توليد جبر الحوادث Generating an Algebra of Events

لقد قدّمنا في الفصل الرابع مفهومي جبر الحوادث و الـ $\sigma$- **جبر** لحوادث مع توضيحات واستنتاجات كثيرة كانت موضع اهتمامنا في ذلك الفصل، ولذلك سنقدّم فيما يلي مبرهنات نقبلها دون برهان. إنّ هذه المبرهنات تُوضّح لنا كيفية توليد جبر الحوادث من حادث أو صف حوادث غير خال $\mathbb{S} \subseteq 2^{\Omega}$، وللوصول إلى هذه الغاية سنعرّف بعض الرموز اللازمة لذلك.

لتكن $\Omega$ مجموعة نتائج تجربة عشوائية (غير خالية)، ولنأخذ $\mathbb{S} \subseteq 2^{\Omega}$ صفًا غير خالٍ من الحوادث، ولنضع بالتعريف ما يلي:

$$\mathbf{C}(\mathbb{S}) = \mathbb{S} \cup \left\{ \overline{A} \; ; A \in \mathbb{S} \right\}$$

$$\mathbf{D}(\mathbb{S}) = \left\{ \bigcap_{k=1}^{n} A_k \;\; ; A_k \in \mathbb{S} \;, k \in \mathbb{N}_n \right\}$$

$$\mathbf{V}(\mathbb{S}) = \left\{ \bigcup_{k=1}^{n} A_k \;\; ; A_k \in \mathbb{S} \;, k \in \mathbb{N}_n \right\}$$

$$\mathbf{V}_d(\mathbb{S}) = \left\{ \bigcup_{k=1}^{n} A_k \; ; A_k \in \mathbb{S} \text{ with } A_k \bigcap_{k \neq \ell} A_\ell = \varnothing \;, k, \ell \in \mathbb{N}_n \right\}$$

#### B-١-١ - مبرهنة

لتكن $\Omega$ مجموعة نتائج تجربة عشوائية (غير خالية)، ولنأخذ $\mathbb{S} \subseteq 2^{\Omega}$ صفًا غير خالٍ من الحوادث، فعندئذ يمكن توليد جبر $\mathscr{A}$ فوق $\Omega$ من خلال إحدى العمليّات الآتية:

**a)** $\mathscr{A} = \mathbf{V}\left(\mathbf{D}\left(\mathbf{C}(\mathbb{S})\right)\right)$

**b)** $\mathscr{A} = \mathbf{V}\left(\mathbf{C}\left(\mathbf{V}\left(\mathbf{C}(\mathbb{S})\right)\right)\right)$

c) $\mathscr{A} = \mathbf{V}_d\left(\mathbf{D}\left(\mathbf{C}(\mathbb{S})\right)\right)$

وحينئذ يُدعى الصف $\mathbb{S}$ **صفاً مولّداً** للجبر $\mathscr{A}$ (**ويُرمز عادةً للجبر المولّد من صف مجموعات $\mathbb{S}$ بـ** $\alpha(\mathbb{S})$).

### B-١-٢- أمثلة

١- إنَّ الجبر المبتذل يولّد على النحو الآتي:

نأخذ الصف $\mathcal{Z} = \{\Omega\}$ (**التجزئة المبتذلة**)، فنجد أنّها تولّد أضعف جبر فوق $\Omega$ لأنّه باستخدام العملية **(a)** نجد ما يلي:

$$\mathscr{C} := \mathbf{C}(\mathcal{Z}) = \mathcal{Z} \cup \{\bar{A} \mid A \in \mathcal{Z}\} = \{\Omega\} \cup \{\varnothing\} = \{\Omega, \varnothing\}$$

$$\mathscr{D} := \mathbf{D}(\mathbf{C}(\mathcal{Z})) = \mathbf{D}(\mathscr{C}) = \left\{\bigcap_{k=1}^{n} A_k \;;\; A_k \in \mathscr{C},\, k \in \mathbb{N}_n\right\} = \{\Omega, \varnothing\}$$

$$\mathscr{A} = \alpha(\mathcal{Z}) = \mathbf{V}\left(\mathbf{D}\left(\mathbf{C}(\mathcal{Z})\right)\right) = \mathbf{V}(\mathscr{D}) = \left\{\bigcup_{k=1}^{n} A_k \;;\; A_k \in \mathscr{D},\, k \in \mathbb{N}_n\right\} = \{\Omega, \varnothing\}$$

كما يمكن استخدام بقية العمليّات للحصول على النتيجة نفسها.

٢- يمكننا توليد جبر من حادث $A \subseteq \Omega$ على النحو الآتي:

نأخذ $A \subseteq \Omega$ حادثاً، ونضع $\mathbb{S} := \{A\}$، فعندئذ يكون لدينا:

$$\mathscr{C} = \mathbf{C}(\mathbb{S}) = \mathbb{S} \cup \{\bar{A} \;;\; A \in \mathbb{S}\} = \{A\} \cup \{\bar{A}\} = \{A, \bar{A}\}$$

$$\mathscr{D} = \mathbf{D}(\mathbf{C}(\mathbb{S})) = \mathbf{D}(\mathscr{C}) = \left\{\bigcap_{k=1}^{n} A_k \;;\; A_k \in \mathscr{C},\, k \in \mathbb{N}_n\right\} = \{A, \bar{A}, \varnothing\}$$

$$\mathscr{A} = \alpha(\mathbb{S}) = \mathbf{V}\left(\mathbf{D}\left(\mathbf{C}(\mathbb{S})\right)\right) = \mathbf{V}(\mathscr{D}) = \left\{\bigcup_{k=1}^{n} A_k \;;\; A_k \in \mathscr{D},\, k \in \mathbb{N}_n\right\} = \{\varnothing, A, \bar{A}, \Omega\}$$

كما يمكن استخدام بقية العمليّات للحصول على النتيجة نفسها.

٣- يمكننا توليد جبر من صف حوادث بسيطة على النحو الآتي:

على سبيل التوضيح والتبسيط سنأخذ المجموعة $\Omega := \{\omega_1, \omega_2, \omega_3\}$، ولنأخذ الصف $\mathbb{S} := \{\{\omega_1\}, \{\omega_2\}\}$، فعندئذ يكون لدينا الآتي:

$$\mathscr{C} = \mathbf{C}(\mathbb{S}) = \mathbb{S} \cup \{\bar{A} \;;\; A \in \mathbb{S}\} = \{\{\omega_1\}, \{\omega_2\}\} \cup \{\{\omega_2, \omega_3\}, \{\omega_1, \omega_3\}\} = \{\{\omega_1\}, \{\omega_2\}, \{\omega_2, \omega_3\}, \{\omega_1, \omega_3\}\}$$

$$\mathscr{D} = \mathbf{D}(\mathbf{C}(\mathbb{S})) = \mathbf{D}(\mathscr{C}) = \left\{\bigcap_{k=1}^{n} A_k \;;\; A_k \in \mathscr{C},\, k \in \mathbb{N}_n\right\} = \{\{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_2, \omega_3\}, \{\omega_1, \omega_3\}, \varnothing\}$$

$$\mathscr{A} = \alpha(\mathbb{S}) = \mathbf{V}\left(\mathbf{D}\left(\mathbf{C}(\mathbb{S})\right)\right) = \mathbf{V}(\mathscr{D}) = \left\{\bigcup_{k=1}^{n} A_k \;;\; A_k \in \mathscr{D},\, k \in \mathbb{N}_n\right\}$$
$$= \{\varnothing, \{\omega_1\}, \{\omega_2\}, \{\omega_3\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}, \{\omega_1, \omega_2, \omega_3\} = \Omega\} = 2^{\Omega}$$

وهو الجبر المولّد من الصف $\mathbb{S}$، حيث يُلاحظ أنّه الجبر القوي (**أقوى جبر**) فوق $\Omega$.

٤- كذلك يمكننا تبيّان أنَّ الجبر القوي $2^{\Omega}$ يكون مولّداً من:

- التجزئة $\mathcal{Z} = \{\{\omega_1\}, \{\omega_2\}, \ldots, \{\omega_n\}\}$ عندما تكون $\Omega$ مجموعة منتهية بقدرة تساوي $|\Omega| = n$.

- التجزئة $\mathscr{Z}=\{\{\omega_1\},\{\omega_2\},\ldots,\{\omega_n\},\ldots\}$ عندما تكون $\Omega$ مجموعة غير منتهية ولكنَّها قابلة للعدّ على الأكثر.

وقبل ختام هذه الفقرة نورد المبرهنة الآتية التي تعرض إحدى طرائق توليد الجبور فوق $\Omega$ أيضاً.

**B-١-٣ - مبرهنة**

ليكن $\mathbb{S}\subset 2^{\Omega}$ صفاً غير خالٍ من الحوادث، فعندئذ صف الحوادث الآتي:

$$\mathscr{A}=\bigcap_{i\in I\,;\,\mathscr{A}_i\supseteq\mathbb{S}}\mathscr{A}_i\quad ;\ \mathscr{A}_i \textit{ is an algebra on } \Omega \textit{ for all } i\in I$$

هو من جديد جبرٌ فوق $\Omega$، وهو أصغر جبر يحوي الصف $\mathbb{S}$، وحينئذ يُدعى الصف $\mathbb{S}$ بالصف المولّد للجبر $\mathscr{A}$، ويُرمز له بـ $\alpha(\mathbb{S})$.

نشير هنا إلى أنَّ عمليّة توليد الـ $\sigma$ - **جبر** ليست بالبساطة التي وجدناها لدى توليد الجبر، إذ إنَّ العمليّات السابقة لتوليد الجبر لا يمكن سحبها على توليد الـ $\sigma$ - **جبر** في الحالة العامّة.

تقدّم لنا المبرهنة الآتية كيفية توليد $\sigma$ - **جبر** من صف حوادث غير خالٍ $\mathbb{S}\subset 2^{\Omega}$.

**B-١-٤ - مبرهنة**

ليكن $\mathbb{S}\subset 2^{\Omega}$ صفاً غير خالٍ من الحوادث، فعندئذ صف الحوادث الآتي:

$$\mathscr{A}=\bigcap_{i\in I\,;\,\mathscr{A}_i\supseteq\mathbb{S}}\mathscr{A}_i\quad ;\ \mathscr{A}_i \textit{ is a } \sigma\textit{-algebra on } \Omega \textit{ for all } i\in I$$

هو من جديد $\sigma$ - **جبر** فوق $\Omega$، وهو أصغر $\sigma$ - **جبر** يحوي الصف $\mathbb{S}$، وفي هذه الحالة يُدعى $\mathbb{S}$ الصف المولّد للـ $\sigma$ - **جبر** $\mathscr{A}$، ويُرمز له بـ $\sigma(\mathbb{S})$.

**B-١-٥ - أمثلة**

١ - لتكن لدينا $\Omega=\{1,2\}$، فعندئذ نجد أنَّ:

أ - صف المجموعات $\mathbb{S}_1:=\{\Omega\}\subset 2^{\Omega}$ يولّد الـ $\sigma$ - **جبر** $\mathscr{A}_1=\{\varnothing,\Omega\}$.

ب - صف المجموعات $\mathbb{S}_2=\{\{1\}\}\subset 2^{\Omega}$ يولّد الـ $\sigma$ - **جبر** $\mathscr{A}_2=\{\varnothing,\{1\},\{2\},\Omega\}$.

ج - صف المجموعات $\mathbb{S}_3=\{\{2\}\}\subset 2^{\Omega}$ يولّد الـ $\sigma$ - **جبر** $\mathscr{A}_3=\{\varnothing,\{1\},\{2\},\Omega\}$.

د - صف المجموعات $2^{\Omega}$ يولّد نفسه، وهو $\sigma$ - **جبر** أيضاً حيث لدينا $\mathscr{A}_4=2^{\Omega}=\{\varnothing,\{1\},\{2\},\{1,2\}=\Omega\}$.

ونتيجةً لما سبق في هذا المثال نلاحظ أنَّ $2^{\Omega}=\mathscr{A}_4\supset\mathscr{A}_3=\mathscr{A}_2\supset\mathscr{A}_1$، وهكذا نلاحظ أنَّه من أجل $\Omega$ مثّبته استطعنا الحصول على أسرة $\sigma$ - **جبور** بحيث تكون هذه الجبور محتواة بعضها في البعض الآخر. بمعنى أنَّ جبر الحوادث المستخدم يمكن أن يولَّد من صفوف عديدة غير خالية من الحوادث، وهذا يعني أنَّ جبر الحوادث المستخدم ليس وحيداً في الحالة العامّة.

**B-١-٦ - الحقل البوريلي Borel-Field**

ليكن $(M\,,\,\rho)$ فضاءً مترياً Metric Space، ولنأخذ صف المجموعات $\mathbb{S}:=\{A\subseteq M\ ;\ A\ \ open\}$، فعندئذ الـ $\sigma$ - **جبر** المولّد بالصف $\mathbb{S}$ (أي $\sigma(\mathbb{S})$) يُسمّى الـ $\sigma$ - **جبر** للمجموعات البوريلية في $M$، وفي الحالة الخاصّة عندما يكون $M=\mathbb{R}^n$ مع $n\in\mathbb{N}$ كيفي ولكن مثّبت، فإنَّ $\sigma(\mathbb{S})$ يُدعى الحقل البوريلي، وذلك لأنَّ $\mathbb{R}^n$ مع دالّة المسافة الإقليدية $\rho(x\,,y)=\|x-y\|$ يُشكّل حقلاً، علماً أنَّ كل مجموعة $A\in\sigma(\mathbb{S})$ تُسمّى مجموعة بوريلية، وقد استخدمنا سابقاً الرمز $\mathfrak{R}^n$ للدلالة على الحقل البوريلي الخاص بـ

$\mathbb{R}^n$ . هذا ويمكن إثبات أنَّ كل مجموعة بوريلية يمكن أن تكتب على شكل اتحاد لعدد غير منته من المجالات المفتوحة لأنَّه من أجل أيّة مجموعة مفتوحة $A$ من $\Re^n$ ستكون حاوية جميع نقاطها الداخلية، ومن ثمَّ يمكننا أنَّ نكتب الآتي:

$$A = \bigcup_{x \in A \; ; \; A \supseteq V_{\varepsilon}(x)} V_{\varepsilon}(x)$$

علماً أنَّ $V_{\varepsilon}(x)$ هي مجاورة لـ $x$ بنصف قطر $\varepsilon>0$، وفي الحالة الخاصّة عندما $n=1$ يمكننا من أجل $A_{k,n} := \left(\frac{k}{n}, \frac{k+1}{n}\right)$ أن نكتب الآتي:

$$A = \bigcup_{n \in \mathbb{N}, k \in \mathbb{Q} \; ; \; A_{k,n} \subseteq A} A_{k,n}$$

**B-١-٧ - مبرهنة**

إنَّ أيّ صف من الصفوف الآتية يُشكّل صفاً مولّداً للـ σ- **جبر** $\Re$ **(الـ σ- جبر لكلّ المجموعات البوريلية في** $\mathbb{R}$**):**

$$\mathcal{S}_1 = \{(\alpha, \beta) \subseteq \mathbb{R} \; ; \; \alpha, \beta \in \mathbb{Q}\} \qquad \& \qquad \mathcal{S}_2 = \{[\alpha, \beta) \subseteq \mathbb{R} \; ; \; \alpha, \beta \in \mathbb{Q}\}$$

$$\mathcal{S}_3 = \{(\alpha, \beta] \subseteq \mathbb{R} \; ; \; \alpha, \beta \in \mathbb{Q}\} \qquad \& \qquad \mathcal{S}_4 = \{[\alpha, \beta] \subseteq \mathbb{R} \; ; \; \alpha, \beta \in \mathbb{Q}\}$$

$$\mathcal{S}_5 = \{(-\infty, \beta] \subseteq \mathbb{R} \; ; \; \beta \in \mathbb{Q}\} \qquad \& \qquad \mathcal{S}_5 = \{(-\infty, \beta) \subseteq \mathbb{R} \; ; \; \beta \in \mathbb{Q}\}$$

$$\mathcal{S}_7 = \{[\alpha, +\infty) \subseteq \mathbb{R} \; ; \; \alpha \in \mathbb{Q}\} \qquad \& \qquad \mathcal{S}_7 = \{(\alpha, +\infty) \subseteq \mathbb{R} \; ; \; \alpha \in \mathbb{Q}\}$$

**B-٢ - الصفوف المطّردة Monotone classes**

ليكن $\mathfrak{M} \subset 2^{\Omega}$ صفاً غير خالٍ، فعندئذ يُقال إنَّ $\mathfrak{M}$ هو صف مطّرد **إذا وفقط إذا** تحقّق ما يلي:

١- من أجل أيّة متتالية مجموعات $(A_n)_{n\in\mathbb{N}}$ غير متناقصة، وعناصرها من $\mathfrak{M}$، فإنَّه سيكون لدينا $\bigcup_{n=1}^{\infty} A_n \in \mathfrak{M}$ أيضاً.

٢- من أجل أيّة متتالية مجموعات $(A_n)_{n\in\mathbb{N}}$ غير متزايدة، وعناصرها من $\mathfrak{M}$، فإنَّه سيكون لدينا $\bigcap_{n=1}^{\infty} A_n \in \mathfrak{M}$ أيضاً.

تعرض لنا المبرهنات الآتية بعض خصائص الصفوف المطّردة والخصائص المتبادلة بين الجبور والصفوف المطّردة.

**B-٢-١ - مبرهنات**

١- إنَّ تقاطع أيّة أسرة غير خالية من الصفوف المطّردة هو من جديد صف مطّرد.

٢- إذا كان $2^{\Omega} \supset \mathcal{S} \neq \varnothing$، فعندئذ الصف الناتج عن تقاطع كل الصفوف المطّردة الحاوية على الصف $\mathcal{S}$ هو من جديد صف مطّرد على $\Omega$. أي إنَّ الصف الآتي:

$$\mathfrak{M} = \bigcap_{i \in I \; ; \; \mathfrak{M}_i \supseteq \mathcal{S}} \mathfrak{M}_i \quad ; \mathfrak{M}_i \textit{ is a mnoton system for all } i \in I$$

هو صف مطّرد على $\Omega$، وهو أصغر صف مطّرد يحوي $\mathcal{S}$، وفي هذه الحالة يُدعى $\mathfrak{M}$ بالصف المولّد للصف المطّرد $\mathcal{S}$، ويُرمز له بـ $\mu(\mathcal{S})$.

٣- إذا كان $\mathscr{A}$ جبراً على $\Omega$، فعندئذ سيكون لدينا ما يلي محقّقاً: $\mu(\mathscr{A}) = \sigma(\mathscr{A})$، علماً أنَّ $\mu(\mathscr{A})$ يرمز إلى الصف المطّرد المولّد من $\mathscr{A}$.

٤- إنَّ كل $\sigma$- **جبر** على $\Omega$ هو صف مطّرد على $\Omega$ أيضاً.

٥- إذا كان $\mathscr{A}$ جبراً على $\Omega$، فعندئذ يكون $\mathscr{A}$ $\sigma$- **جبراً** فوق $\Omega$ **إذا وفقط إذا** كان $\mathscr{A}$ صفاً مطّرداً.

### B-٣ - التطبيقات القيوسة Measurable maps

إنَّ التطبيقات القيوسة تُعدّ من المفاهيم المهمّة جداً في دراسة نظرية القياس عامّة وفي مجال نظرية الاحتمالات خاصّة ذلك أنَّ هذا المفهوم يُعدّ الحجر الأساس لمفهوم العناصر العشوائية في نظرية الاحتمالات، ولهذا لا بدَّ للقارئ من الفهم الجيد لهذا الجانب.

سنقوم فيما يلي بتقديم بسيط وموجز حول التطبيقات القيوسة ومن ثمَّ سنعود لاحقاً (بعد تقديم القياس) إلى تناول الدوال القيوسة Measurable Functions التي تُعدّ حالة خاصّة من التطبيقات القيوسة.

لتكن $\Omega \neq \varnothing$ و $G \neq \varnothing$ مجموعتين كيفيتين، وبفرض أنَّ $\mathscr{A}$ و $\mathscr{G}$ هما $\sigma$- **جبر** فوق $\Omega$ و $G$ على الترتيب، ولنأخذ $X : \Omega \to G$ تطبيقاً مفترضاً، وسنضع بالتعريف ما يلي:

$$X^{-1}(B) := \{\omega \in \Omega \,;\, X(\omega) \in B\} \quad ; B \in \mathscr{G}$$

$$X^{-1}(\mathscr{G}) := \left\{X^{-1}(B) \,;\, B \in \mathscr{G}\right\} \subseteq \mathscr{A}$$

عندئذ ينتج لدينا من خواص الصورة العكسية للتطبيقات ما يلي:

$$X^{-1}(\varnothing) = \varnothing$$

$$\overline{X^{-1}(B)} = X^{-1}(\bar{B}) \quad ; B \in \mathscr{G}$$

$$X^{-1}\left(\bigcup_{i \in I} B_i\right) = \bigcup_{i \in I} X^{-1}(B_i) \quad ; B_i \in \mathscr{G}, \forall i \in I$$

$$X^{-1}\left(\bigcap_{i \in I} B_i\right) = \bigcap_{i \in I} X^{-1}(B_i) \quad ; B_i \in \mathscr{G}, \forall i \in I$$

#### B-٣-١ - تعريف التطبيق القيوس:

بفرض أنَّ $\mathscr{A}$ و $\mathscr{G}$ هما $\sigma$- **جبر** فوق $\Omega$ و $G$ على الترتيب، وليكن $X : \Omega \to G$ تطبيقاً مُعطى، فإذا كان من أجل كل $B \in \mathscr{G}$ لدينا $X^{-1}(B)$ هو عنصر من $\mathscr{A}$، فعندئذ يُقال إنَّ التطبيق $X$ **قيوسٌ بالنسبة إلى** $\mathscr{A}$ و $\mathscr{G}$.

على سبيل المثال لو أخذنا $\Omega = \{1,2\}$ مجموعة مُعطاة، ولنأخذ الـ $\sigma$- **جبر** $\mathscr{A} = \{\varnothing, \Omega\}$ (الجبر الضعيف فوق $\Omega$)، فعندئذ بفرض أنَّ $X$ تطبيقٌ حقيقي مُعرَّفٌ على $\Omega$ من خلال العلاقة $X(\omega) = 0$، فإنَّنا سنجد ما يلي:

$$\{X(\omega) < x\} = \begin{cases} \varnothing & for \ \ x \leq 0 \\ \Omega & for \ \ x > 0 \end{cases}$$

وبما أنّه لدينا بحسب الفرض $\varnothing , \Omega \in \mathscr{A}$، فإنَّ ذلك يعني أنَّ التطبيق $X$ قيوساً بالنسبة إلى $\mathscr{A}$ و $\Re$.

#### B-٣-٢ - بعض خصائص التطبيقات القيوسة:

١- تركيب تطبيقين قيوسين هو من جديد تطبيق قيوس.

٢- إذا كانت $A$ و $B \in \mathscr{G}$ مع $A \bigcap B = \varnothing$، فعندئذ إذا كان $X$ تطبيقاً قيوساً بالنسبة إلى $\mathscr{A}$ و $\mathscr{G}$ فإنَّه سيكون لدينا:

$$X^{-1}(A) \bigcap X^{-1}(B) = \varnothing \in \mathscr{A}$$

٣- بفرض أنَّ $\mathscr{A}$ هو $\sigma$- **جبر** فوق $\Omega$، وكذلك لدينا $\mathscr{G}$ هو $\sigma$- **جبر** فوق $G$، وبفرض أنَّ $X$ تطبيقاً قيوساً بالنسبة إلى $\mathscr{A}$

و $\mathcal{G}$، فعندئذ سيكون الصف $X^{-1}(\mathcal{G})$ هو $\sigma$ **- جبر** فوق $\Omega$ أيضاً، وهو أصغر $\sigma$ **- جبر** فوق $\Omega$ يكون من أجله التطبيق $X$ قيوساً بالنسبة إلى $\mathcal{A}$ و $\mathcal{G}$.

**٤-** بفرض أنَّ $\mathcal{A}$ و $\mathcal{G}$ هما $\sigma$ **- جبر** فوق $\Omega$ و $G$ على الترتيب، وبفرض أنَّ $\mathcal{S}$ صف مجموعات مُولِّداً للـ $\sigma$ **- جبر** $\mathcal{G}$، ولنأخذ $X : \Omega \to G$ تطبيقاً مُعطى، فعندئذ يكون التطبيق $X$ قيوساً بالنسبة إلى $\mathcal{A}$ و $\mathcal{G}$ **إذا وفقط إذا** كان $X^{-1}(\mathcal{S}) \subset \mathcal{A}$.

**٥-** بفرض أنَّ $\mathcal{A}$ هو $\sigma$ **- جبر** فوق $\Omega$، وكذلك $\mathcal{G}_i$ هو $\sigma$ **- جبر** فوق $G_i$ من أجل كل $i \ni I$، و $I$ مجموعة منتهية أو غير منتهية ولكن قابلة للعدّ على الأكثر، ولنأخذ أسرة تطبيقات قيوسة $\{X_i \;; i \in I\}$ بحيث إن التطبيق $X_i$ قيوس بالنسبة إلى $\mathcal{A}$ و $\mathcal{G}_i$ من أجل كل $i \ni I$. عندئذ يكون الاتحاد $\bigcup_{i \in I} X_i^{-1}(\mathcal{G}_i)$ هو $\sigma$ **- جبر** فوق $\Omega$، وهو أصغر $\sigma$ **- جبر** تكون من أجله التطبيقات $X_i$ قيوسة من أجل كل $i \ni I$.

### B-٣-٣ - تعريف ( الـ $\sigma$ - جبر المولَّد من تطبيق قيوس):

بفرض أنَّ $\mathcal{A}$ و $\mathcal{G}$ هما $\sigma$ **- جبر** فوق $\Omega$ و $G$ على الترتيب، ولنأخذ $X$ تطبيقاً قيوساً بالنسبة إلى $\mathcal{A}$ و $\mathcal{G}$، فعندئذ $X^{-1}(\mathcal{G})$ يُدعى **الـ $\sigma$ - جبر المولَّد بالتطبيق** $X$، ويُرمز له بـ $\sigma(X)$، أي إنَّه لدينا:

$$\sigma(X) = X^{-1}(\mathcal{G}) = \left\{X^{-1}(B) \;; \forall B \in \mathcal{G}\right\}$$

فنلاحظ هنا أنَّ الـ $\sigma$ **- جبر** المولَّد بالتطبيق $X$ هو $\sigma$ **- جبر** جزئي من $\mathcal{A}$، وذلك لأنَّ $X^{-1}(\mathcal{G}) \subseteq \mathcal{A}$.

### B-٣-٣ - تعريف (الـ $\sigma$ - جبر المولَّد من أسرة تطبيقات قيوسة):

بفرض أنَّ $\mathcal{A}$ هو $\sigma$ **- جبر** فوق $\Omega$، و $\mathcal{G}_i$ هو $\sigma$ **- جبر** فوق $G_i$ مع $i \ni I$ علماً أنَّ $I$ مجموعة منتهية أو غير منتهية ولكن قابلة للعدّ على الأكثر، ولنأخذ أسرة تطبيقات قيوسة $\{X_i \;; i \in I\}$ بحيث يكون التطبيق $X_i$ قيوساً بالنسبة إلى $\mathcal{A}$ و $\mathcal{G}_i$ من أجل كل $i \ni I$. عندئذ صف المجموعات النّاتج عن $\bigcup_{i \in I} X_i^{-1}(\mathcal{G}_i)$ يولِّد $\sigma$ **- جبر** فوق $\Omega$، ويُدعى **الـ $\sigma$ - جبر المولَّد بأسرة التطبيقات القيوسة** $\{X_i \;; i \in I\}$، ويُرمز له بـ $\sigma(X_i \;; i \in I)$، أي إنَّ:

$$\sigma(X_i \;; i \in I) = \sigma\left(\bigcup_{i \in I} X^{-1}(\mathcal{G}_i)\right)$$

## B-٤ - فضاءات القياس Measure Spaces:

لقد لاحظنا أنَّ الفضاء الاحتمالي يُعدّ من المفاهيم المهمَّة جداً في نظرية الاحتمالات لأنَّه يكوِّن الركيزة الأساس في نمذجة التجارب العشوائية، ولكن هذا المفهوم ما هو إلاً حالة خاصّة من مفهوم فضاء القياس، حيث يُعدّ هذا المفهوم الأخير من المفاهيم المهمَّة جداً في نظرية القياس أيضاً وذلك لأنَّه بمثابة الأساس الذي تقوم عليه دراسة التطبيقات القيوسة وتكاملاتها.

### B-٤-١ - تعريف (الفضاء المقيس):

بفرض أنَّ $\mathcal{A}$ هو $\sigma$ **- جبر** فوق $\Omega$، فعندئذ تُدعى الثنائية $[\Omega , \mathcal{A}]$ **فضاءً مقيساً**.

### B-٤-٢ - تعريف (القياس Measure):

ليكن $[\Omega , \mathcal{A}]$ فضاءً مقيساً، و $\mu : \mathcal{A} \to [0, \infty]$ دالَّة (**دالَّة مجموعات**) مُعطاة، فإذا حقَّقت الدالَّة $\mu$ الخاصتين الآتيتين:

١- لدينا $\mu(\varnothing) = 0$ (**هذا يعني أنَّ $\mu$ ليست مبتذلة، بمعنى أنَّه يوجد على الأقل $A \ni \mathcal{A}$ بحيث يكون** $\mu(A) > 0$).

٢- من أجل أيّة متتالية مجموعات $(A_n)_{n \in \mathbb{N}}$ عناصرها من $\mathcal{A}$ ومنفصلة مثنى مثنى يكون لدينا:

$$\mu\left(\sum_{n=1}^{\infty} A_n\right)=\sum_{n=1}^{\infty}\mu(A_n)$$

فعندئذ يُقال عن الدالّة $\mu$ **إنّها قياس** على الفضاء المقيس $[\Omega , \mathscr{A}]$ .

**B-٤-٣ - حالات خاصة:**

ليكن $\mu$ قياساً على الفضاء المقيس $[\Omega , \mathscr{A}]$، فعندئذ:

١- إذا كان لدينا $\mu(\Omega)<+\infty$ ، فعندئذ يُدعى $\mu$ **قياساً منتهياً** Finite Measure.

٢- إذا كان لدينا $\mu(\Omega)=1$، فعندئذ يُدعى $\mu$ **قياساً احتمالياً**.

٣- إذا كان $\mu(\Omega)=+\infty$، وأمكننا إيجاد متتالية مجموعات $(Z_n)_{n\in\mathbb{N}}$ عناصرها من $\mathscr{A}$ تشكّل تجزئة للمجموعة $\Omega$، وبحيث إنّه من أجل كل $n\in\mathbb{N}$ لدينا $\mu(Z_n)<+\infty$ ، فعندئذ يُقال إنّ $\mu$ هو قياس من **نوع** $\sigma-$ **منتهٍ** σ-Finite Measure، وبعض المراجع العربية تذكره باسم **قياس منتهٍ عدود**.

**B-٤-٤ - خصائص القياس:**

ليكن $\mu$ قياساً على فضاء مقيس $[\Omega , \mathscr{A}]$، فعندئذ:

١- من أجل أيّة مجموعتين $A,B\in\mathscr{A}$ مع $A\subseteq B$ يكون لدينا $\mu(A)\leq\mu(B)$ ، وهذه الخاصيّة تعرف باسم خاصيّة **الاطّراد** Monotony للقياس $\mu$ .

٢- من أجل أيّة مجموعتين $A,B\in\mathscr{A}$ مع $A\subseteq B$ وكذلك $\mu(B)<+\infty$ تكون العلاقة الآتية مُحقّقة:

$$\mu(B\setminus A)=\mu(B)-\mu(A)$$

٣- من أجل أيّة مجموعتين $A,B\in\mathscr{A}$ يكون لدينا:

$$\mu(A\cup B)\ \leq\ \mu(A)+\mu(B)$$

وهذه الخاصيّة يُعبَّر عنها بالقول: إنَّ القياس $\mu$ **نصف جمعي** Subadditive.

٤- من أجل أيّة مجموعات $A_1,A_2,\ldots,A_n\in\mathscr{A}$ يكون لدينا:

$$\mu\left(\bigcup_{i=1}^{n} A_i\right)\leq\sum_{i=1}^{n}\mu(A_i)$$

وهذه الخاصيّة يُعبَّر عنها بالقول: إنَّ القياس $\mu$ **نصف جمعي منتهٍ** Finite Subadditive.

٥- من أجل أيّة متتالية مجموعات $(A_n)_{n\in\mathbb{N}}$ عناصرها من $\mathscr{A}$ يكون لدينا:

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right)\leq\sum_{n=1}^{\infty}\mu(A_n)$$

وهذه الخاصيّة يُعبَّر عنها بالقول: إنَّ القياس $\mu$ **نصف جمعي** $\sigma-$ **منتهٍ** σ-Finite Subadditive (أو نصف جمعي عدود).

نشير أخيراً إلى ضرورة التمييز بين تعبير القياس الوارد في الفصل الأول والقياس الذي يستخدم في نظرية القياس رغم أنّه يُستخدم من أجلهما الكلمة Measure نفسها.

### B-٤-٥ - تعريف (قياس لوبيغ Lebesgue Measure)

لنأخذ الفضاء المقيس $[\mathbb{R},\mathfrak{R}]$، وليكن $\lambda$ دالّة مجموعات على معرَّفة $\mathfrak{R}$ كما يلي:

$$\lambda([a,b)) = b-a \qquad ; a,b \in \mathbb{R}\ , b>a$$

فعندئذ نجد أنَّ $\lambda$ تُحقِّق شروط القياس على $[\mathbb{R},\mathfrak{R}]$، وذلك لأنَّه:

١- لدينا $\lambda(\varnothing)=\lambda([a\ ,a))=a-a=0$ من أجل أي $a \in \mathbb{R}$.

٢- من أجل أيّة متتالية فترات $\left([a_n\ ,b_n)\right)_{n\in N}$ عناصرها من $\mathfrak{R}$ ومنفصلة مثنى مثنى، سيكون لدينا ما يلي مُحَقَّقاً:

$$\lambda\left(\sum_{n=1}^{\infty}[a_n\ ,b_n)\right)=\sum_{n=1}^{\infty}\lambda([a_n\ ,b_n))$$

### B-٤-٦ - ملاحظات

١- نشير هنا إلى أنَّه من أجل قياس لوبيغ يوضع من أجل كل $a, b \in \mathbb{R}$ مع $b>a$ بالتعريف ما يلي:

$$\lambda([a,b))=\lambda((a\ ,b\,])=\lambda([a,b\,])=b-a$$

٢- إضافة لما سبق فإنَّ $\lambda$ هو قياس من نوع $\sigma$- منتهٍ وذلك لأنَّه لو أخذ التجزئة $\mathcal{Z}=\{\,[n\ ,n+1);\ n\in \mathbb{Z}\,\}$ (وهي مجموعة جزئية من $\mathfrak{R}$) للمجموعة $\mathbb{R}$ ، فإنَّنا نجد ما يلي:

$$\textbf{a)}\quad \lambda\left([n\ ,n+1)\right)=(n+1)-n=1<+\infty$$

$$\textbf{b)}\quad \lambda(\mathbb{R})=\lambda\left(\sum_{n=-\infty}^{\infty}[n\ ,n+1)\right)=\sum_{n=-\infty}^{\infty}\lambda([n\ ,n+1))=\sum_{n=-\infty}^{\infty}[(n+1)-n]=\sum_{n=-\infty}^{\infty}1=+\infty$$

٣- إنَّ قياس لوبيغ على الفضاء المقيس $\left[\mathbb{R}^n\ ,\mathfrak{R}^n\right]$ الذي يُرمز له بـ $\lambda^n$ ، يُعرَّف بشكل مماثل لما سبق حيث يوضع بالتعريف:

$$\lambda^n\left(\underset{i=1}{\overset{n}{\times}}[a_i\ ,\ b_i)\right)=\prod_{i=1}^{n}(b_i-a_i) \qquad ; a_i\ ,b_i\in\mathbb{R}\ ,b_i>a_i$$

٤- يُبرهن على أنَّ قياس لوبيغ $\lambda^n$ المُعطى سابقاً وحيداً على الفضاء المقيس $\left[\mathbb{R}^n\ ,\mathfrak{R}^n\right]$. إنَّ معنى وحدانية قياس لوبيغ على $\left[\mathbb{R}^n\ ,\mathfrak{R}^n\right]$ سيتضّح لنا لاحقاً لدى تقديم مبرهنة رادون- نيكوديم.

### B-٤-٧ - الاستمرار من الأدنى والأعلى لقياس Lower and Upper Continuity of a measure

ليكن $\mu$ دالّة مجموعات على فضاء مقيس $[\Omega\ ,\mathcal{A}]$ ، فعندئذ يُقال:

١- إنَّ الدالّة $\mu$ مستمرَّة من الأدنى إذا كان من أجل أية متتالية مجموعات $\left(A_n\right)_{n\in N}$ عناصر من $\mathcal{A}$ وغير متناقصة، فإنَّ العلاقة الآتية تكون مُحَقَّقةً:

$$\mu\left(\lim_{n\to\infty}A_n\right)=\mu\left(\bigcup_{n=1}^{\infty}A_n\right)=\lim_{n\to\infty}\mu(A_n)$$

٢- إنَّ الدالّة $\mu$ مستمرَّة من الأعلى إذا كان من أجل أية متتالية مجموعات $\left(A_n\right)_{n\in N}$ عناصر من $\mathcal{A}$ وغير متزايدة، فإنَّ العلاقة الآتية تكون مُحَقَّقةً:

$$\mu\left(\lim_{n\to\infty}A_n\right)=\mu\left(\bigcap_{n=1}^{\infty}A_n\right)=\lim_{n\to\infty}\mu(A_n)$$

المبرهنة الآتية تبيِّن لنا أنَّ القياس مستمرّ من الأدنى والأعلى.

### B-٤-٨ - مبرهنة

ليكن $\mu$ قياساً على فضاء مقيس $[\Omega , \mathscr{A}]$. عندئذ يكون $\mu$ مستمرّاً من الأدنى والأعلى بآن واحد (ولذلك يُقال إنَّ القياس $\mu$ مستمر).

### B-٤-٩ - ملاحظات

١- إذا كان $\mu(\Omega) < +\infty$ فإنَّه من أجل أي $A \in \mathscr{A}$ تكون العلاقة الآتية مُحَقَّقةً:

$$\mu(\bar{A}) = \mu(\Omega \setminus A) = \mu(\Omega) - \mu(A)$$

٢- إذا كانت $\mu_1$ ، $\mu_2$ ، ...، $\mu_n$ و... متتالية قياسات منتهية فوق فضاء مقيس $[\Omega , \mathscr{A}]$ ، و $(a_n)_{n\in\mathbb{N}}$ متتالية عددية من $\mathbb{R}^+$ ، فعندئذ تكون دالَّة المجموعات:

$$\mu := \sum_{n=1}^{\infty} a_n \cdot \mu_n$$

قياساً على $[\Omega , \mathscr{A}]$ أيضاً، وعلاوةً على ذلك:

أ- يكون القياس $\mu$ منته إذا وفقط إذا كان:

$$\sum_{n=1}^{\infty} a_n \cdot \mu_n(\Omega) < +\infty \,.$$

ب- إذا كانت القياسات $\mu_1$ ، $\mu_2$ ، ...، $\mu_n$ و... هي قياسات احتمالية، فعندئذ يكون $\mu$ قياساً احتمالياً إذا وفقط إذا كان:

$$\sum_{n=1}^{\infty} a_n = 1$$

المبرهنة الآتية تعرض لنا الصيغة التي تؤول إليها القياسات المنتهية عندما تكون $\Omega$ مجموعة قابلة للعدّ.

### B-٤-١٠ - مبرهنة

لتكن $\Omega$ مجموعة قابلة للعدّ على الأكثر، ولنأخذ $\mathscr{A} = 2^{\Omega}$ ، وليكن $\mu$ قياسًا من النوع $\sigma$- منته على $[\Omega , \mathscr{A}]$، فعندئذ يكون للقياس $\mu$ العرض الآتي:

$$\mu(A) = \sum_{\omega \in \Omega} \mu(\{\omega\}) \cdot \delta_{\omega}(A) \quad ; \forall A \in \mathscr{A}$$

علماً أنَّ $\delta_{\omega}$ يرمز إلى قياس ديراك المُركَّز على النقطة $\omega$ .

### B-٤-١١ - تعريف (فضاء القياس Measure Space)

ليكن $\mu$ قياساً على فضاء مقيس $[\Omega , \mathscr{A}]$، فعندئذ تُدعى الثلاثية $[\Omega , \mathscr{A}, \mu]$ فضاء قياس.

في الحالة الخاصّة إذا كان $\mu$ قياساً احتمالياً، فعندئذ يُقال عن فضاء القياس $[\Omega , \mathscr{A}, \mu]$ إنَّه فضاء احتمالي.

### B-٤-١٢ - تعريف (القياس المولَّد من تطبيق قيوس)

ليكن $[\Omega , \mathscr{A}, \mu]$ فضاءً قياساً و $[G , \mathscr{G}]$ فضاءً مقيساً، ولنأخذ $X : \Omega \to G$ تطبيقاً قيوساً بالنسبة إلى $\mathscr{A}$ و $\mathscr{G}$، فعندئذ يُسمَّى القياس $\eta := \mu \circ X^{-1}$ قياساً مولَّداً بالتطبيق $X$ ، ويُرمز عادةً لهذا القياس بـ $\mu_X$، وهكذا يمكننا من أجل أيّ $B \in \mathscr{G}$ أن نكتب الآتي:

$$\eta(B)=\mu_X(B)=\left(\mu\circ X^{-1}\right)(B)=\mu\left(X^{-1}(B)\right)$$
$$=\mu\{\{\omega\in\Omega\ ;X(\omega)\in B\}\}\triangleq\mu(X\in B)$$

### B-٤-١٣- حالات خاصّة

١- إذا كان القياس $\mu$ منتهياً، فإنَّ القياس $\mu_X$ سيكون منتهياً أيضاً.

٢- إذا كان القياس $\mu$ من النوع $\sigma$- منته، فإنَّ القياس $\mu_X$ سيكون من النوع $\sigma$- منته أيضاً.

٣- إذا كان القياس $\mu$ احتمالياً، فإنَّ القياس $\mu_X$ سيكون قياساً احتمالياً أيضاً، ومن أجل هذه الحالة الخاصّة غالباً يستخدم الرمز $P$ و $P_X$ عوضاً عن $\mu$ و $\mu_X$ على الترتيب، ومن ثمَّ يمكننا أن نكتب من أجل أيّ $B\in\mathfrak{B}$ الآتي:

$$P_X(B)=P\circ X^{-1}(B)=P\left(X^{-1}(B)\right)=\int_{X^{-1}(B)}P(d\omega)=\int\mathbf{I}_{X^{-1}(B)}(\omega)\,P(d\omega)$$
$$=\int\mathbf{I}_B\left(X(\omega)\right)P(d\omega)=\int\mathbf{I}_B(x)P_X(dx)$$

وذلك لأنَّ:

$$P_X(B)=\int_B P_X(dx)=\int P_X(dx)\,\mathbf{I}_B(x)$$

تبحث المبرهنة الآتية في وجود القياس الناتج عن جداء عددمنته من القياسات.

### B-٤-١٤- مبرهنة

لتكن $\mu_1$ ، $\mu_2$ ، ...، $\mu_n$ قياسات من النوع $\sigma$- منته على الفضاءات المقيسة $[\Omega_1\ ,\mathscr{A}_1]$، $[\Omega_2\ ,\mathscr{A}_2]$، ...و $[\Omega_n\ ,\mathscr{A}_n]$ على الترتيب، ولنأخذ $\bigotimes_{i=1}^{n}\mathscr{A}_i$ هو الـ $\sigma$- **جبر** فوق $\bigtimes_{i=1}^{n}\Omega_i$ . عندئذ **يوجد** قياس وحيد $\mu$ يُرمز له بـ $\mu:=\mu_1\times\mu_1\times\ldots\times\mu_n$ على الفضاء المقيس $\left[\bigtimes_{i=1}^{n}\Omega_i\ ,\bigotimes_{i=1}^{n}\mathscr{A}_i\right]$، وبحيث إنَّه من أجل أيّة مجموعات $A_1\in\mathscr{A}_1$ ، $A_2\in\mathscr{A}_2$ ،...و $A_n\in\mathscr{A}_n$ تكون العلاقة الآتية مُحقَّقة:

$$\mu\left(\bigtimes_{k=1}^{n}A_k\right)=\prod_{k=1}^{n}\mu_k(A_k)$$

وهو قياس من النوع $\sigma$- منته.

بناءً على المبرهنة السابقة يكون قياس لوبيغ $\lambda^n$ هو قياس ضرب (جداء) على $[\mathbb{R}^n\ ,\mathfrak{R}^n]$ ، ومن ثمَّ يمكن كتابته على شكل ضرب لقياس لوبيغ بنفسه $n$ مرّة أيضاً. أي إنَّه يمكننا أن نكتب $\lambda^n=\lambda\times\lambda\times\ldots\times\lambda$ .

## B-٥- الدوال القيوسة Measurable Functions

سنقدِّم فيما يلي مفهوم الدالَّة القيوسة التي تُعدّ حالة خاصّة من التطبيق القيوس.

### B-٥-١- تعريف (الدالَّة القيوسة)

لتكن $f$ دالَّة حقيقية معرَّفة على $\Omega$ ، فعندئذٍ يُقال عن $f$ إنَّها **دالَّة قيوسة** (أو دالَّة قابلة للقياس) **إذا وفقط إذا** كانت قيوسة بالنسبة إلى $\mathscr{A}$ و $\mathfrak{R}$ ، وعادة يختصر التعبير إلى القول إنَّ $f$ قيوسة بالنسبة إلى $\mathscr{A}$ دون ذكر $\mathfrak{R}$ في هذه الحالة الخاصّة (أي إذا كانت $f$ دالَّة حقيقية).

تقدِّم المبرهنة الآتية بعض الخصائص للدوال القيوسة.

**B-٥-٢- مبرهنة**

١- بفرض أنَّ $f$ و $g$ دالَّتين حقيقيتين معرَّفتين على $\Omega$ وقيوستين بالنسبة إلى $\mathscr{A}$، فعندئذ:

أ- من أجل أيّ $a, b \in \mathbb{R}$ مع $a \neq 0$ تكون الدالَّة $af + b$ قيوسة بالنسبة إلى $\mathscr{A}$ أيضاً.

ب- تكون الدالَّة $|f|$ قيوسة بالنسبة إلى $\mathscr{A}$ أيضاً.

ج- تكون الدالَّة $f \pm g$ قيوسة بالنسبة إلى $\mathscr{A}$ أيضاً.

د- تكون الدالَّة $f \cdot g$ قيوسة بالنسبة إلى $\mathscr{A}$ أيضاً.

٢- بفرض أنَّ $(f_n)_{n\in\mathbb{N}}$ متتالية دوال حقيقية قيوسة بالنسبة إلى $\mathscr{A}$، فعندئذ ستكون الدوال الآتية (**في حال وجودها**) قيوسة بالنسبة إلى $\mathscr{A}$ أيضاً:

$$f = \inf\{f_n \ ; n \in \mathbb{N}\} \quad \& \quad f = \sup\{f_n \ ; n \in \mathbb{N}\} \quad \& \quad f = \overline{\lim_{n\to\infty}} f_n \quad \& \quad f = \underline{\lim_{n\to\infty}} f_n \quad \& \quad f = \lim_{n\to\infty} f_n$$

**B-٥-٣- ملاحظة**

إنَّ الاختبارات التي قُدِّمت من أجل التحقّق من قيوسيَّة التطبيقات تبقى سارية المفعول من أجل الدوال الحقيقية أيضاً.

**B-٥-٤- تعريف (تكامل دالَّة حقيقية قيوسة Integral of a Measurable Function)**

ليكن $[\Omega, \mathscr{A}, \mu]$ فضاءً قياساً، ولنأخذ $f$ دالَّة حقيقية معرَّفة على $\Omega$ وقيوسة بالنسبة إلى $\mathscr{A}$. عندئذ يُرمز لتكامل الدالَّة $f$ بالنسبة إلى القياس $\mu$ بأحد العروض الآتية:

$$\mathbf{E}_\mu f \quad \text{or} \quad \int f \, d\mu \quad \text{or} \quad \int f(\omega)\, \mu(d\omega)$$

ويُعرَّف على النحو الآتي:

١- إذا كانت الدالَّة $f$ بسيطةً أو درجية، أي إنَّ $f = \sum_{i=1}^{n} a_i \cdot \mathbf{I}_{A_i}$، فإنَّ تكامل الدالَّة $f$ بالنسبة إلى القياس $\mu$ يُحسب بالعلاقة الآتية:

$$\mathbf{E}_\mu f = \sum_{i=1}^{n} a_i \cdot \mu(A_i)$$

٢- إذا كانت الدالَّة $f$ غير سالبة، فعندئذ توجد متتالية دوال بسيطة $(f_n)_{n\in\mathbb{N}}$ بحيث يكون $\lim_{n\to\infty} f_n = f$، ومن ثم تكامل $f$ بالنسبة إلى $\mu$ يُحسب بالعلاقة الآتية:

$$\mathbf{E}_\mu f = \lim_{n\to\infty} \int f_n \, d\mu$$

٣- إذا كانت الدالَّة $f$ كيفية، فعندئذ يُمكننا أن نكتب $f = f^+ - f^-$، علماً أنَّ:

$$f^+(\omega) = \max\{f(\omega), 0\} = \begin{cases} 0 & for\ f(\omega) < 0 \\ f(\omega) & for\ f(\omega) \geq 0 \end{cases}$$

$$f^{-}(\omega) = -\min\{f(\omega), 0\} = \begin{cases} 0 & for \quad f(\omega) \geq 0 \\ -f(\omega) & for \quad f(\omega) < 0 \end{cases}$$

فإذا كان $\int f^{-} d\mu < +\infty$ أو $\int f^{+} d\mu < +\infty$ ، فإنَّ تكامل الدالَّة $f$ بالنسبة إلى $\mu$ يُعطى بالعلاقة الآتية:

$$\mathbf{E}_{\mu} f = \int f^{+} d\mu - \int f^{-} d\mu$$

### B-٥-٥- تعريف (قابلية المكاملة لدالَّة حقيقية Integrablity of a Real Function)

لتكن $f : \Omega \to \overline{\mathbb{R}}$ دالَّة قيوسة بالنسبة إلى $\mathscr{A}$ ، فإذا كان $\int f \, d\mu < +\infty$ ، فعندئذ يُقال إنَّ $f$ **دالَّة كمولة** (**أو قابلة للمكاملة**) بالنسبة إلى القياس $\mu$ .

في الحقيقة يمكن البرهنة على المقولات (المبرهنات) الآتية:

١- إذا كانت $f : \Omega \to \overline{\mathbb{R}}$ دالَّة قيوسة بالنسبة إلى $\mathscr{A}$ ، فعندئذ تكون الدالَّة $f$ كمولة بالنسبة إلى القياس $\mu$ **إذا وفقط إذا** كانت العلاقة $\int |f| \, d\mu < +\infty$ محقَّقةً.

٢- إذا كانت $f : [a, b] \to \overline{\mathbb{R}}$ قيوسة بالنسبة إلى $\Re \cap [a, b]$ ، وكمولة بحسب مفهوم ريمان، فعندئذ يكون لدينا:

$$\mathbf{E}_{\lambda} f = \int_a^b f(x) \ \lambda(dx) = \int_a^b f(x) \ dx$$

٣- بفرض أنَّ $[\Omega, \mathscr{A}, \mu]$ فضاء قياس مُعطى، و $f : \Omega \to \overline{\mathbb{R}}$ دالَّة غير سالبة وقيوسة بالنسبة إلى $\mathscr{A}$ ، وبفرض أنَّ $(A_n)_{n \in \mathbb{N}}$ متتالية مجموعات عناصرها من $\mathscr{A}$ ، ولنضع:

$$A := \sum_{n=1}^{\infty} A_n$$

فعندئذ إذا كانت $f$ كمولة بالنسبة إلى $\mu$ فإنَّ العلاقة الآتية ستكون محقَّقة:

$$\int_A f \ d\mu = \int_{\sum_{n=1}^{\infty} A_n} f \ d\mu = \sum_{n=1}^{\infty} \int_{A_n} f \ d\mu$$

### B-٥-٦- نظرية بيبو- ليفي Beppo Levi's Theorem

لتكن $(f_n)_{n \in \mathbb{N}}$ متتالية دوال حقيقية غير سالبة، وليست متناقصة على $\mathbb{R}$ ، وقيوسة بالنسبة إلى قياس $\sigma$- منته $\mu$ على $[\mathbb{R}, \Re]$ ، فعندئذ إذا كانت المتتالية $(f_n)_{n \in \mathbb{N}}$ متقاربة من دالَّة حقيقية $f$ (تقارباً نقطياً على الأقل)، فإنَّ العلاقة الآتية ستكون مُحقَّقة:

$$\lim_{n \to \infty} \int f_n \ d\mu = \int \lim_{n \to \infty} f_n \ d\mu = \int f \ d\mu$$

وتعرف هذه المبرهنة باسم **نظرية التقارب المطَّرد** Monotone Convergence Theoremأيضاً .

### B-٥-٦- نظرية فوبيني Fubini Theorem

ليكن $\mu_1$ و $\mu_2$ قياسين من نوع $\sigma$- منته، ولنأخذ $f : \Omega_1 \times \Omega_2 \to \overline{\mathbb{R}}$ دالَّة حقيقية وقيوسة بالنسبة إلى $\mathscr{A}_1 \otimes \mathscr{A}_2$ ، فإذا كان التكامل $\int f(\omega_1, \omega_2) \ \mu_1 \times \mu_2 \, d(\omega_1, \omega_2)$ موجوداً، فعندئذ سيكون كل من التكاملين الآتيين موجوداً أيضاً:

$$\int f(\omega_1, \omega_2) \ \mu_1(d\omega_1) \qquad ; \forall \ \omega_2 \in \Omega$$

$$\int f\left(\omega_1, \omega_2\right) \mu_2\left(d\, \omega_2\right) \qquad ; \forall\, \omega_1 \in \Omega$$

وتكون العلاقة الآتية مُحَقَّقة:

$$\int f\left(\omega_1, \omega_2\right) \mu_1 \times \mu_2\; d\left(\omega_1, \omega_2\right) = \int \left[\, \int f\left(\omega_1, \omega_2\right)\, \mu_1(d\, \omega_1)\right] \mu_2(d\, \omega_2)$$
$$= \int \left[\, \int f\left(\omega_1, \omega_2\right) \mu_2(d\, \omega_2)\right] \mu_1(d\, \omega_1)$$

إنَّ هذه النظرية تعود إلى الرياضياتي الإيطالي **فوبيني** (1879 –1943) Guido Fubini.

### B-٥-٧- تعريف (يساوي تقريباً في كل مكان Equal Almost Every Where)

ليكن $[\Omega\,,\mathscr{A}, \mu]$ فضاءً قياساً، ولنأخذ $f$ و $g$ دالَّتين حقيقيتين غير سالبتين معرفتين على $\Omega$، فعندئذٍ يُقال إنَّ $f$ **تساوي** $g$ **تقريباً في كل مكان** بالنسبة إلى القياس $\mu$ **إذا وفقط إذا** كانت العلاقة الآتية مُحَقَّقة:

$$\mu\left(\left\{\omega \in \Omega\ ; f\ (\omega) \neq g\ (\omega)\right\}\right) = 0$$

وتُكتب هذه العلاقة بأحد الأشكال الآتية:

$$f \underset{\mu}{\sim} g \quad \text{or} \quad f \underset{\mu}{=} g \quad \text{or} \quad f = g \ \ ; \mu\text{–}a.e.$$

علماً أنَّ العبارة $f = g\ ; \mu\text{–}a.e.$ (**يساوي تقريباً في كل مكان**) دارجة الاستخدام في التحليل الرياضياتي وقليل ما تستخدم في العشوائيات، ولكن إذا كان $\mu$ قياساً احتمالياً، فعندئذٍ يُقال إنَّ $f = g$ **باحتمال يساوي الواحد** أو **باحتمال شبه أكيد**، ويُكتب ذلك على النحو الاتي:

$$f = g \ \ ; \mu\text{–}a.s.$$

وكذلك تجدر الإشارة هنا إلى أنَّ العلاقة $\sim$ السابقة هي في الحقيقة علاقة تكافؤ بين $f$ و $g$، ولذلك يُقال في حال $f \underset{\mu}{\sim} g$ إنَّ الدالَّتين $f$ و $g$ متكافئتان وفقاً للقياس $\mu$.

أخيراً نشير إلى أنَّ التكافؤ:

$$f \underset{\mu}{=} \vartheta \Leftrightarrow \int f \ \ d\mu = 0$$

مُحقَّق دوماً، علماً أنَّ $\vartheta$ ترمز إلى الدالَّة الصفرية على $\Omega$.

### B-٥-٨- تعريف (الاستمرار المطلق Absolute Continuation)

ليكن $\mu$ و $\eta$ قياسين على فضاء مقيس $[\Omega\,, \mathscr{A}]$. عندئذٍ يُقال إنَّ القياس $\eta$ **مستمرٌ مطلقاً** بالنسبة إلى القياس $\mu$ (**ويُرمز لذلك بـ** $\eta \prec\prec \mu$) **إذا وفقط إذا** كان من أجل أي $A \in \mathscr{A}$ مع $\mu(A) = 0$ فإنَّه ينتج لدينا $\eta(A) = 0$ أيضاً.

### B-٥-٩- نظرية رادون – نيكوديم Radon–Nikodym Theorem

ليكن $\mu$ و $\eta$ قياسين على فضاء مقيس $[\Omega\,, \mathscr{A}]$، فإذا كان $\mu$ من نوع $\sigma$– منته، فعندئذ تكون المقولتان الآتيتان متكافئتين:

١- توجد دالَّة حقيقية غير سالبة $f$ على $\Omega$ منتهية تقريباً في كل مكان بالنسبة إلى $\mu$ مع:

$$\eta(A) = \int_A f\ (\omega)\ \ \mu(d\omega) \qquad ; \forall A \in \mathscr{A}$$

٢- القياس $\eta$ هو قياس $\sigma$– منته ومستمر مطلقاً بالنسبة إلى القياس $\mu$.

### B-٥-١٠ ملاحظات

١- إنَّ الدالَّة $f$ المذكورة في نص المبرهنة السابقة تُسمَّى **مشتق رادون – نيكوديم** أو **الكثافة من القياس** $\eta$ **بالنسبة إلى القياس** $\mu$ ، ويُكتب ذلك بالشكل الآتي:

$$\frac{d\,\eta}{d\,\mu} = f$$

٢-علاوةً على ما سبق فإنَّ الدالَّة $f$ **معيَّنة بشكل وحيد**. أي إنَّه إذا وجدت دالَّة أخرى غير سالبة $g$ على $\Omega$ مع:

$$\eta(A) = \int_A g(\omega)\,\mu(d\omega) \quad ;\forall A \in \mathfrak{F} \;\Rightarrow\; \mu\big(\{\omega \in \Omega\,; f(\omega) \neq g(\omega)\}\big) = 0$$

٣- إنَّ الدالَّة $f$ الواردة في نص المبرهنة السابقة كمولة بالنسبة إلى القياس $\mu$ **إذا وفقط إذا** كان القياس $\eta$ منتهٍ

٤- لقد ذكرنا سابقاً لدى عرض مبرهنة الوحدانية لقياس لوبيغ على $\mathbb{R}^n$ أنَّنا سنذكر لاحقاً معنى هذه الوحدانية. في الواقع إنَّ معنى الوحدانية في تلك المبرهنة يعني أنَّه من أجل أي قياس آخر $\eta$ معرَّف على $\mathbb{R}^n$ يجب أن يكون هذا القياس مستمراً مطلقاً بالنسبة إلى قياس لوبيغ $\lambda^n$ ، ومن ثمَّ وجود دالَّة غير سالبة $f$ قيوسة وكمولةٍ بالنسبة إلى $\lambda^n$ من أجلها يكون لدينا ما يلي مُحَقَّقاً:

$$\frac{d\,\eta}{d\,\lambda^n} = f \;\Leftrightarrow\; d\,\eta = f\;d\,\lambda^n \;\Leftrightarrow\; \int_A d\,\eta = \int_A f\;d\,\lambda^n \quad ;\forall A \in \Re^n$$

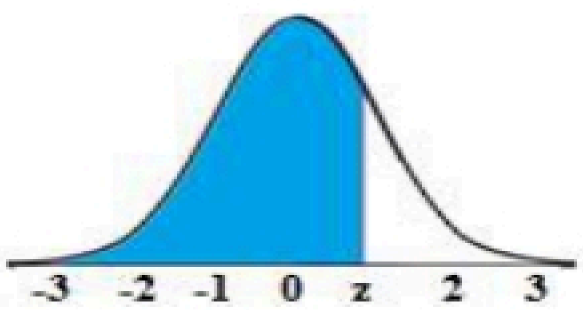
**بهذا نختم الملحق *B***

# الملحق (C)

# جداول إحصائية

# وملخص

# لتوزيعات احتمالية شهيرة



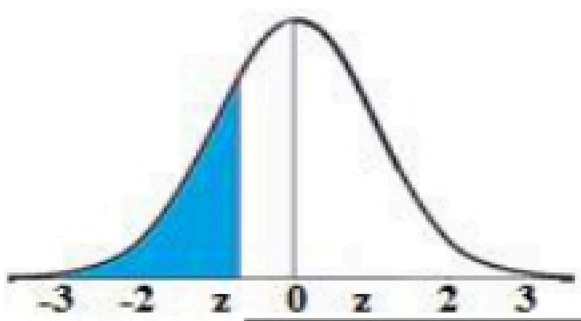

## جدول قيم التوزيع الطبيعي المعياري

## Standard Normal Distribution Table

من أجل القيم الموجبة لـ Z

| z | 0.00 | 0.01 | 0.02 | 0.03 | 0.04 | 0.05 | 0.06 | 0.07 | 0.08 | 0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| 0.0 | 0.5000 | 0.5040 | 0.5080 | 0.5120 | 0.5160 | 0.5199 | 0.5239 | 0.5279 | 0.5319 | 0.5359 |
| 0.1 | 0.5398 | 0.5438 | 0.5478 | 0.5517 | 0.5557 | 0.5596 | 0.5636 | 0.5675 | 0.5714 | 0.5753 |
| 0.2 | 0.5793 | 0.5832 | 0.5871 | 0.5910 | 0.5948 | 0.5987 | 0.6026 | 0.6064 | 0.6103 | 0.6141 |
| 0.3 | 0.6179 | 0.6217 | 0.6255 | 0.6293 | 0.6331 | 0.6368 | 0.6406 | 0.6443 | 0.6480 | 0.6517 |
| 0.4 | 0.6554 | 0.6591 | 0.6628 | 0.6664 | 0.6700 | 0.6736 | 0.6772 | 0.6808 | 0.6844 | 0.6879 |
| 0.5 | 0.6915 | 0.6950 | 0.6985 | 0.7019 | 0.7054 | 0.7088 | 0.7123 | 0.7157 | 0.7190 | 0.7224 |
| 0.6 | 0.7257 | 0.7291 | 0.7324 | 0.7357 | 0.7389 | 0.7422 | 0.7454 | 0.7486 | 0.7517 | 0.7549 |
| 0.7 | 0.7580 | 0.7611 | 0.7642 | 0.7673 | 0.7704 | 0.7734 | 0.7764 | 0.7794 | 0.7823 | 0.7852 |
| 0.8 | 0.7881 | 0.7910 | 0.7939 | 0.7967 | 0.7995 | 0.8023 | 0.8051 | 0.8078 | 0.8106 | 0.8133 |
| 0.9 | 0.8159 | 0.8186 | 0.8212 | 0.8238 | 0.8264 | 0.8289 | 0.8315 | 0.8340 | 0.8365 | 0.8389 |
| 1.0 | 0.8413 | 0.8438 | 0.8461 | 0.8485 | 0.8508 | 0.8531 | 0.8554 | 0.8577 | 0.8599 | 0.8621 |
| 1.1 | 0.8643 | 0.8665 | 0.8686 | 0.8708 | 0.8729 | 0.8749 | 0.8770 | 0.8790 | 0.8810 | 0.8830 |
| 1.2 | 0.8849 | 0.8869 | 0.8888 | 0.8907 | 0.8925 | 0.8944 | 0.8962 | 0.8980 | 0.8997 | 0.9015 |
| 1.3 | 0.9032 | 0.9049 | 0.9066 | 0.9082 | 0.9099 | 0.9115 | 0.9131 | 0.9147 | 0.9162 | 0.9177 |
| 1.4 | 0.9192 | 0.9207 | 0.9222 | 0.9236 | 0.9251 | 0.9265 | 0.9279 | 0.9292 | 0.9306 | 0.9319 |
| 1.5 | 0.9332 | 0.9345 | 0.9357 | 0.9370 | 0.9382 | 0.9394 | 0.9406 | 0.9418 | 0.9429 | 0.9441 |
| 1.6 | 0.9452 | 0.9463 | 0.9474 | 0.9484 | 0.9495 | 0.9505 | 0.9515 | 0.9525 | 0.9535 | 0.9545 |
| 1.7 | 0.9554 | 0.9564 | 0.9573 | 0.9582 | 0.9591 | 0.9599 | 0.9608 | 0.9616 | 0.9625 | 0.9633 |
| 1.8 | 0.9641 | 0.9649 | 0.9656 | 0.9664 | 0.9671 | 0.9678 | 0.9686 | 0.9693 | 0.9699 | 0.9706 |
| 1.9 | 0.9713 | 0.9719 | 0.9726 | 0.9732 | 0.9738 | 0.9744 | 0.9750 | 0.9756 | 0.9761 | 0.9767 |
| 2.0 | 0.9772 | 0.9778 | 0.9783 | 0.9788 | 0.9793 | 0.9798 | 0.9803 | 0.9808 | 0.9812 | 0.9817 |
| 2.1 | 0.9821 | 0.9826 | 0.9830 | 0.9834 | 0.9838 | 0.9842 | 0.9846 | 0.9850 | 0.9854 | 0.9857 |
| 2.2 | 0.9861 | 0.9864 | 0.9868 | 0.9871 | 0.9875 | 0.9878 | 0.9881 | 0.9884 | 0.9887 | 0.9890 |
| 2.3 | 0.9893 | 0.9896 | 0.9898 | 0.9901 | 0.9904 | 0.9906 | 0.9909 | 0.9911 | 0.9913 | 0.9916 |
| 2.4 | 0.9918 | 0.9920 | 0.9922 | 0.9925 | 0.9927 | 0.9929 | 0.9931 | 0.9932 | 0.9934 | 0.9936 |
| 2.5 | 0.9938 | 0.9940 | 0.9941 | 0.9943 | 0.9945 | 0.9946 | 0.9948 | 0.9949 | 0.9951 | 0.9952 |
| 2.6 | 0.9953 | 0.9955 | 0.9956 | 0.9957 | 0.9959 | 0.9960 | 0.9961 | 0.9962 | 0.9963 | 0.9964 |
| 2.7 | 0.9965 | 0.9966 | 0.9967 | 0.9968 | 0.9969 | 0.9970 | 0.9971 | 0.9972 | 0.9973 | 0.9974 |
| 2.8 | 0.9974 | 0.9975 | 0.9976 | 0.9977 | 0.9977 | 0.9978 | 0.9979 | 0.9979 | 0.9980 | 0.9981 |
| 2.9 | 0.9981 | 0.9982 | 0.9982 | 0.9983 | 0.9984 | 0.9984 | 0.9985 | 0.9985 | 0.9986 | 0.9986 |
| 3.0 | 0.9987 | 0.9987 | 0.9987 | 0.9988 | 0.9988 | 0.9989 | 0.9989 | 0.9989 | 0.9990 | 0.9990 |
| 3.1 | 0.9990 | 0.9991 | 0.9991 | 0.9991 | 0.9992 | 0.9992 | 0.9992 | 0.9992 | 0.9993 | 0.9993 |
| 3.2 | 0.9993 | 0.9993 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9994 | 0.9995 | 0.9995 | 0.9995 |
| 3.3 | 0.9995 | 0.9995 | 0.9995 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9996 | 0.9997 |
| 3.4 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9997 | 0.9998 |

يتبع ⇐

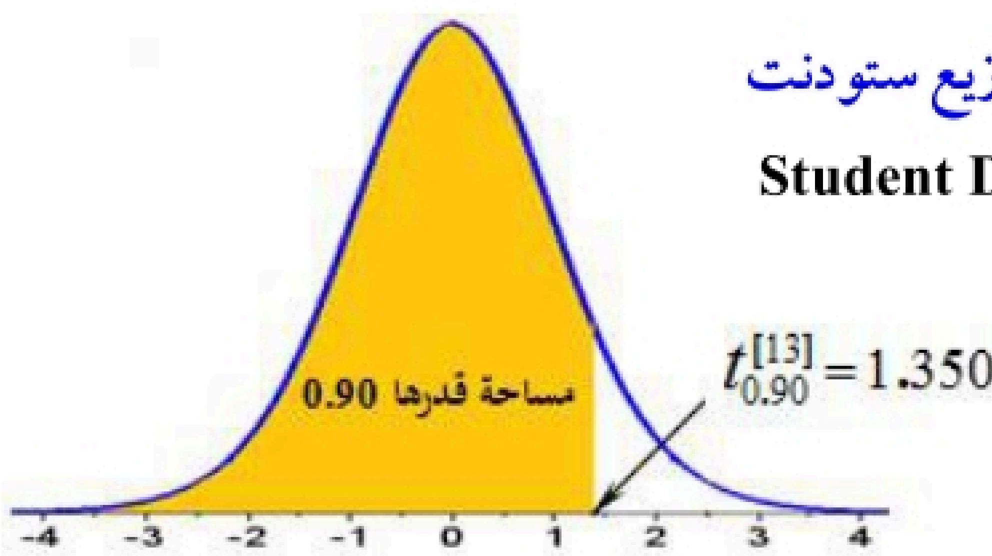



## تابع جدول قيم التوزيع الطبيعي المعياري

## Standard Normal Distribution Table

من أجل القيم السالبة لـ Z

| z | 0.00 | -0.01 | -0.02 | -0.03 | -0.04 | -0.05 | -0.06 | -0.07 | -0.08 | -0.09 |
|---|---|---|---|---|---|---|---|---|---|---|
| - 3.4 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0003 | 0.0002 |
| - 3.3 | 0.0005 | 0.0005 | 0.0005 | 0.0004 | 0.0004 | 0.0004 | 0.0004 | 0.0004 | 0.0004 | 0.0003 |
| - 3.2 | 0.0007 | 0.0007 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0006 | 0.0005 | 0.0005 | 0.0005 |
| - 3.1 | 0.0010 | 0.0009 | 0.0009 | 0.0009 | 0.0008 | 0.0008 | 0.0008 | 0.0008 | 0.0007 | 0.0007 |
| - 3.0 | 0.0013 | 0.0013 | 0.0013 | 0.0012 | 0.0012 | 0.0011 | 0.0011 | 0.0011 | 0.0010 | 0.0010 |
| - 2.9 | 0.0019 | 0.0018 | 0.0018 | 0.0017 | 0.0016 | 0.0016 | 0.0015 | 0.0015 | 0.0014 | 0.0014 |
| - 2.8 | 0.0026 | 0.0025 | 0.0024 | 0.0023 | 0.0023 | 0.0022 | 0.0021 | 0.0021 | 0.0020 | 0.0019 |
| - 2.7 | 0.0035 | 0.0034 | 0.0033 | 0.0032 | 0.0031 | 0.0030 | 0.0029 | 0.0028 | 0.0027 | 0.0026 |
| - 2.6 | 0.0047 | 0.0045 | 0.0044 | 0.0043 | 0.0041 | 0.0040 | 0.0039 | 0.0038 | 0.0037 | 0.0036 |
| - 2.5 | 0.0062 | 0.0060 | 0.0059 | 0.0057 | 0.0055 | 0.0054 | 0.0052 | 0.0051 | 0.0049 | 0.0048 |
| - 2.4 | 0.0082 | 0.0080 | 0.0078 | 0.0075 | 0.0073 | 0.0071 | 0.0069 | 0.0068 | 0.0066 | 0.0064 |
| - 2.3 | 0.0107 | 0.0104 | 0.0102 | 0.0099 | 0.0096 | 0.0094 | 0.0091 | 0.0089 | 0.0087 | 0.0084 |
| - 2.2 | 0.0139 | 0.0136 | 0.0132 | 0.0129 | 0.0125 | 0.0122 | 0.0119 | 0.0116 | 0.0113 | 0.0110 |
| - 2.1 | 0.0179 | 0.0174 | 0.0170 | 0.0166 | 0.0162 | 0.0158 | 0.0154 | 0.0150 | 0.0146 | 0.0143 |
| - 2.0 | 0.0228 | 0.0222 | 0.0217 | 0.0212 | 0.0207 | 0.0202 | 0.0197 | 0.0192 | 0.0188 | 0.0183 |
| - 1.9 | 0.0287 | 0.0281 | 0.0274 | 0.0268 | 0.0262 | 0.0256 | 0.0250 | 0.0244 | 0.0239 | 0.0233 |
| - 1.8 | 0.0359 | 0.0351 | 0.0344 | 0.0336 | 0.0329 | 0.0322 | 0.0314 | 0.0307 | 0.0301 | 0.0294 |
| - 1.7 | 0.0446 | 0.0436 | 0.0427 | 0.0418 | 0.0409 | 0.0401 | 0.0392 | 0.0384 | 0.0375 | 0.0367 |
| - 1.6 | 0.0548 | 0.0537 | 0.0526 | 0.0516 | 0.0505 | 0.0495 | 0.0485 | 0.0475 | 0.0465 | 0.0455 |
| - 1.5 | 0.0668 | 0.0655 | 0.0643 | 0.0630 | 0.0618 | 0.0606 | 0.0594 | 0.0582 | 0.0571 | 0.0559 |
| - 1.4 | 0.0808 | 0.0793 | 0.0778 | 0.0764 | 0.0749 | 0.0735 | 0.0721 | 0.0708 | 0.0694 | 0.0681 |
| - 1.3 | 0.0968 | 0.0951 | 0.0934 | 0.0918 | 0.0901 | 0.0885 | 0.0869 | 0.0853 | 0.0838 | 0.0823 |
| - 1.2 | 0.1151 | 0.1131 | 0.1112 | 0.1093 | 0.1075 | 0.1056 | 0.1038 | 0.1020 | 0.1003 | 0.0985 |
| - 1.1 | 0.1357 | 0.1335 | 0.1314 | 0.1292 | 0.1271 | 0.1251 | 0.1230 | 0.1210 | 0.1190 | 0.1170 |
| - 1.0 | 0.1587 | 0.1562 | 0.1539 | 0.1515 | 0.1492 | 0.1469 | 0.1446 | 0.1423 | 0.1401 | 0.1379 |
| - 0.9 | 0.1841 | 0.1814 | 0.1788 | 0.1762 | 0.1736 | 0.1711 | 0.1685 | 0.1660 | 0.1635 | 0.1611 |
| - 0.8 | 0.2119 | 0.2090 | 0.2061 | 0.2033 | 0.2005 | 0.1977 | 0.1949 | 0.1922 | 0.1894 | 0.1867 |
| - 0.7 | 0.2420 | 0.2389 | 0.2358 | 0.2327 | 0.2296 | 0.2266 | 0.2236 | 0.2206 | 0.2177 | 0.2148 |
| - 0.6 | 0.2743 | 0.2709 | 0.2676 | 0.2643 | 0.2611 | 0.2578 | 0.2546 | 0.2514 | 0.2483 | 0.2451 |
| - 0.5 | 0.3085 | 0.3050 | 0.3015 | 0.2981 | 0.2946 | 0.2912 | 0.2877 | 0.2843 | 0.2810 | 0.2776 |
| - 0.4 | 0.3446 | 0.3409 | 0.3372 | 0.3336 | 0.3300 | 0.3264 | 0.3228 | 0.3192 | 0.3156 | 0.3121 |
| - 0.3 | 0.3821 | 0.3783 | 0.3745 | 0.3707 | 0.3669 | 0.3632 | 0.3594 | 0.3557 | 0.3520 | 0.3483 |
| - 0.2 | 0.4207 | 0.4168 | 0.4129 | 0.4090 | 0.4052 | 0.4013 | 0.3974 | 0.3936 | 0.3897 | 0.3859 |
| - 0.1 | 0.4602 | 0.4562 | 0.4522 | 0.4483 | 0.4443 | 0.4404 | 0.4364 | 0.4325 | 0.4286 | 0.4247 |
| 0.0 | 0.5000 | 0.4960 | 0.4920 | 0.4880 | 0.4840 | 0.4801 | 0.4761 | 0.4721 | 0.4681 | 0.4641 |




## جدول القيم الحرجة لتوزيع ستودنت

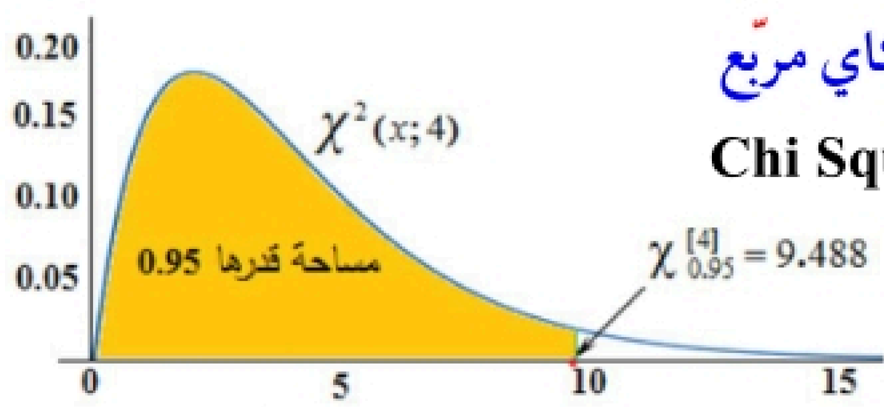
**Student Distribution Table**

| d.f. $n$ | $\alpha$ 0.90 | 0.95 | 0.975 | 0.99 | 0.995 | 0.999 |
|---|---|---|---|---|---|---|
| 1 | 3.078 | 6.314 | 12.706 | 31.821 | 63.657 | 318.313 |
| 2 | 1.886 | 2.920 | 4.303 | 6.965 | 9.925 | 22.327 |
| 3 | 1.638 | 2.353 | 3.182 | 4.541 | 5.841 | 10.215 |
| 4 | 1.533 | 2.132 | 2.776 | 3.747 | 4.604 | 7.173 |
| 5 | 1.476 | 2.015 | 2.571 | 3.365 | 4.032 | 5.893 |
| 6 | 1.440 | 1.943 | 2.447 | 3.143 | 3.707 | 5.208 |
| 7 | 1.415 | 1.895 | 2.365 | 2.998 | 3.499 | 4.782 |
| 8 | 1.397 | 1.860 | 2.306 | 2.896 | 3.355 | 4.499 |
| 9 | 1.383 | 1.833 | 2.262 | 2.821 | 3.250 | 4.296 |
| 10 | 1.372 | 1.812 | 2.228 | 2.764 | 3.169 | 4.143 |
| 11 | 1.363 | 1.796 | 2.201 | 2.718 | 3.106 | 4.024 |
| 12 | 1.356 | 1.782 | 2.179 | 2.681 | 3.055 | 3.929 |
| 13 | 1.350 | 1.771 | 2.160 | 2.650 | 3.012 | 3.852 |
| 14 | 1.345 | 1.761 | 2.145 | 2.624 | 2.977 | 3.787 |
| 15 | 1.341 | 1.753 | 2.131 | 2.602 | 2.947 | 3.733 |
| 16 | 1.337 | 1.746 | 2.120 | 2.583 | 2.921 | 3.686 |
| 17 | 1.333 | 1.740 | 2.110 | 2.567 | 2.898 | 3.646 |
| 18 | 1.330 | 1.734 | 2.101 | 2.552 | 2.878 | 3.610 |
| 19 | 1.328 | 1.729 | 2.093 | 2.539 | 2.861 | 3.579 |
| 20 | 1.325 | 1.725 | 2.086 | 2.528 | 2.845 | 3.552 |
| 21 | 1.323 | 1.721 | 2.080 | 2.518 | 2.831 | 3.527 |
| 22 | 1.321 | 1.717 | 2.074 | 2.508 | 2.819 | 3.505 |
| 23 | 1.319 | 1.714 | 2.069 | 2.500 | 2.807 | 3.485 |
| 24 | 1.318 | 1.711 | 2.064 | 2.492 | 2.797 | 3.467 |
| 25 | 1.316 | 1.708 | 2.060 | 2.485 | 2.787 | 3.450 |
| 26 | 1.315 | 1.706 | 2.056 | 2.479 | 2.779 | 3.435 |
| 27 | 1.314 | 1.703 | 2.052 | 2.473 | 2.771 | 3.421 |
| 28 | 1.313 | 1.701 | 2.048 | 2.467 | 2.763 | 3.408 |
| 29 | 1.311 | 1.699 | 2.045 | 2.462 | 2.756 | 3.396 |
| 30 | 1.310 | 1.697 | 2.042 | 2.457 | 2.750 | 3.385 |
| 31 | 1.309 | 1.696 | 2.040 | 2.453 | 2.744 | 3.375 |
| 32 | 1.309 | 1.694 | 2.037 | 2.449 | 2.738 | 3.365 |
| 33 | 1.308 | 1.692 | 2.035 | 2.445 | 2.733 | 3.356 |
| 34 | 1.307 | 1.691 | 2.032 | 2.441 | 2.728 | 3.348 |
| 35 | 1.306 | 1.690 | 2.030 | 2.438 | 2.724 | 3.340 |
| 36 | 1.306 | 1.688 | 2.028 | 2.434 | 2.719 | 3.333 |
| 37 | 1.305 | 1.687 | 2.026 | 2.431 | 2.715 | 3.326 |
| 38 | 1.304 | 1.686 | 2.024 | 2.429 | 2.712 | 3.319 |
| 39 | 1.304 | 1.685 | 2.023 | 2.426 | 2.708 | 3.313 |
| 40 | 1.303 | 1.684 | 2.021 | 2.423 | 2.704 | 3.307 |
| 41 | 1.303 | 1.683 | 2.020 | 2.421 | 2.701 | 3.301 |
| 42 | 1.302 | 1.682 | 2.018 | 2.418 | 2.698 | 3.296 |
| 43 | 1.302 | 1.681 | 2.017 | 2.416 | 2.695 | 3.291 |
| 44 | 1.301 | 1.680 | 2.015 | 2.414 | 2.692 | 3.286 |
| 45 | 1.301 | 1.679 | 2.014 | 2.412 | 2.690 | 3.281 |
| 46 | 1.300 | 1.679 | 2.013 | 2.410 | 2.687 | 3.277 |
| 47 | 1.300 | 1.678 | 2.012 | 2.408 | 2.685 | 3.273 |
| 48 | 1.299 | 1.677 | 2.011 | 2.407 | 2.682 | 3.269 |
| 49 | 1.299 | 1.677 | 2.010 | 2.405 | 2.680 | 3.265 |
| 50 | 1.299 | 1.676 | 2.009 | 2.403 | 2.678 | 3.261 |

# جدول القيم الحرجة لتوزيع ستودنت

## Student Distribution Table

| d.f. $n$ | $\alpha$ 0.90 | 0.95 | 0.975 | 0.99 | 0.995 | 0.999 |
|---|---|---|---|---|---|---|
| 51 | 1.298 | 1.675 | 2.008 | 2.402 | 2.676 | 3.258 |
| 52 | 1.298 | 1.675 | 2.007 | 2.400 | 2.674 | 3.255 |
| 53 | 1.298 | 1.674 | 2.006 | 2.399 | 2.672 | 3.251 |
| 54 | 1.297 | 1.674 | 2.005 | 2.397 | 2.670 | 3.248 |
| 55 | 1.297 | 1.673 | 2.004 | 2.396 | 2.668 | 3.245 |
| 56 | 1.297 | 1.673 | 2.003 | 2.395 | 2.667 | 3.242 |
| 57 | 1.297 | 1.672 | 2.002 | 2.394 | 2.665 | 3.239 |
| 58 | 1.296 | 1.672 | 2.002 | 2.392 | 2.663 | 3.237 |
| 59 | 1.296 | 1.671 | 2.001 | 2.391 | 2.662 | 3.234 |
| 60 | 1.296 | 1.671 | 2.000 | 2.390 | 2.660 | 3.232 |
| 61 | 1.296 | 1.670 | 2.000 | 2.389 | 2.659 | 3.229 |
| 62 | 1.295 | 1.670 | 1.999 | 2.388 | 2.657 | 3.227 |
| 63 | 1.295 | 1.669 | 1.998 | 2.387 | 2.656 | 3.225 |
| 64 | 1.295 | 1.669 | 1.998 | 2.386 | 2.655 | 3.223 |
| 65 | 1.295 | 1.669 | 1.997 | 2.385 | 2.654 | 3.220 |
| 66 | 1.295 | 1.668 | 1.997 | 2.384 | 2.652 | 3.218 |
| 67 | 1.294 | 1.668 | 1.996 | 2.383 | 2.651 | 3.216 |
| 68 | 1.294 | 1.668 | 1.995 | 2.382 | 2.650 | 3.214 |
| 69 | 1.294 | 1.667 | 1.995 | 2.382 | 2.649 | 3.213 |
| 70 | 1.294 | 1.667 | 1.994 | 2.381 | 2.648 | 3.211 |
| 71 | 1.294 | 1.667 | 1.994 | 2.380 | 2.647 | 3.209 |
| 72 | 1.293 | 1.666 | 1.993 | 2.379 | 2.646 | 3.207 |
| 73 | 1.293 | 1.666 | 1.993 | 2.379 | 2.645 | 3.206 |
| 74 | 1.293 | 1.666 | 1.993 | 2.378 | 2.644 | 3.204 |
| 75 | 1.293 | 1.665 | 1.992 | 2.377 | 2.643 | 3.202 |
| 76 | 1.293 | 1.665 | 1.992 | 2.376 | 2.642 | 3.201 |
| 77 | 1.293 | 1.665 | 1.991 | 2.376 | 2.641 | 3.199 |
| 78 | 1.292 | 1.665 | 1.991 | 2.375 | 2.640 | 3.198 |
| 79 | 1.292 | 1.664 | 1.990 | 2.374 | 2.640 | 3.197 |
| 80 | 1.292 | 1.664 | 1.990 | 2.374 | 2.639 | 3.195 |
| 81 | 1.292 | 1.664 | 1.990 | 2.373 | 2.638 | 3.194 |
| 82 | 1.292 | 1.664 | 1.989 | 2.373 | 2.637 | 3.193 |
| 83 | 1.292 | 1.663 | 1.989 | 2.372 | 2.636 | 3.191 |
| 84 | 1.292 | 1.663 | 1.989 | 2.372 | 2.636 | 3.190 |
| 85 | 1.292 | 1.663 | 1.988 | 2.371 | 2.635 | 3.189 |
| 86 | 1.291 | 1.663 | 1.988 | 2.370 | 2.634 | 3.188 |
| 87 | 1.291 | 1.663 | 1.988 | 2.370 | 2.634 | 3.187 |
| 88 | 1.291 | 1.662 | 1.987 | 2.369 | 2.633 | 3.185 |
| 89 | 1.291 | 1.662 | 1.987 | 2.369 | 2.632 | 3.184 |
| 90 | 1.291 | 1.662 | 1.987 | 2.368 | 2.632 | 3.183 |
| 91 | 1.291 | 1.662 | 1.986 | 2.368 | 2.631 | 3.182 |
| 92 | 1.291 | 1.662 | 1.986 | 2.368 | 2.630 | 3.181 |
| 93 | 1.291 | 1.661 | 1.986 | 2.367 | 2.630 | 3.180 |
| 94 | 1.291 | 1.661 | 1.986 | 2.367 | 2.629 | 3.179 |
| 95 | 1.291 | 1.661 | 1.985 | 2.366 | 2.629 | 3.178 |
| 96 | 1.290 | 1.661 | 1.985 | 2.366 | 2.628 | 3.177 |
| 97 | 1.290 | 1.661 | 1.985 | 2.365 | 2.627 | 3.176 |
| 98 | 1.290 | 1.661 | 1.984 | 2.365 | 2.627 | 3.175 |
| 99 | 1.290 | 1.660 | 1.984 | 2.365 | 2.626 | 3.175 |
| 100 | 1.290 | 1.660 | 1.984 | 2.364 | 2.626 | 3.174 |
| $\infty$ | 1.28 | 1.64 | 1.96 | 2.33 | 2.58 | 3.15 |




# جدول القيم الحرجة لتوزيع كاي مربّع

## Chi Square Distribution Table

| d.f. *n* | α 0.90 | 0.950 | 0.975 | 0.990 | 0.995 | 0.999 |
|---|---|---|---|---|---|---|
| 1 | 2.706 | 3.841 | 5.024 | 6.635 | 7.879 | 10.8276 |
| 2 | 4.605 | 5.991 | 7.378 | 9.210 | 10.597 | 13.8155 |
| 3 | 6.251 | 7.815 | 9.348 | 11.345 | 12.833 | 16.2662 |
| 4 | 7.779 | 9.488 | 11.143 | 13.277 | 14.860 | 18.4668 |
| 5 | 9.236 | 11.070 | 12.832 | 15.086 | 16,750 | 20.5150 |
| 6 | 10.645 | 12.592 | 14.449 | 16.812 | 18.548 | 22.4577 |
| 7 | 12,017 | 14.067 | 16.013 | 18.475 | 20.278 | 24.3219 |
| 8 | 13.362 | 15.507 | 17.535 | 20.090 | 21.955 | 26.1245 |
| 9 | 14.684 | 16.919 | 19.023 | 21.666 | 23.589 | 27.8772 |
| 10 | 15-987 | 18.307 | 20.483 | 23.209 | 25.188 | 29.5883 |
| 11 | 17.275 | 19,675 | 21.920 | 24,725 | 26.757 | 31.2641 |
| 12 | 18,549 | 21.026 | 23.336 | 26.217 | 28.300 | 32.9095 |
| 13 | 19.812 | 22.362 | 24.736 | 27.688 | 29.819 | 34.5282 |
| 14 | 21.064 | 23.685 | 26.119 | 29.141 | 31.319 | 36.1233 |
| 15 | 22.307 | 24.996 | 27.488 | 30.578 | 32.801 | 37.6973 |
| 16 | 23.542 | 26.296 | 28.845 | 32.000 | 34.267 | 39.2524 |
| 17 | 24.769 | 27.587 | 30.191 | 33.409 | 35.718 | 40.7902 |
| 18 | 25.989 | 28.869 | 31.526 | 34.805 | 37.156 | 42.3124 |
| 19 | 27.204 | 30.144 | 32.852 | 36.191 | 38.582 | 43.8202 |
| 20 | 28.412 | 31.410 | 34.170 | 37.566 | 39.997 | 45.3147 |
| 21 | 29.615 | 32-671 | 35.479 | 38.932 | 41.401 | 46.7970 |
| 22 | 30,813 | 33.924 | 36.781 | 40.289 | 42.796 | 48.2679 |
| 23 | 32.007 | 35.172 | 38.076 | 41.638 | 44.181 | 49.7282 |
| 24 | 33.196 | 36.4)5 | 39.364 | 42.980 | 45.558 | 51.1786 |
| 25 | 34.382 | 37.652 | 40.646 | 44.314 | 46.928 | 52.6197 |
| 26 | 35.563 | 38.8S5 | 41.923 | 45.642 | 48-290 | 54.0520 |
| 27 | 36.741 | 40.113 | 43.194 | 46.963 | 49.645 | 55.4760 |
| 28 | 37.916 | 41.337 | 44.461 | 48.278 | 50.993 | 56.8923 |
| 29 | 39.087 | 42.557 | 45.722 | 49.588 | 52.336 | 58.3012 |
| 30 | 40.256 | 43.773 | 46.979 | 50.892 | 53.672 | 59.7031 |
| 35 | 46.059 | 49.802 | 53.203 | 57.342 | 60.275 | 66.6188 |
| 40 | 51.805 | 55.758 | 59.342 | 63.691 | 66.766 | 73.4019 |
| 45 | 57.505 | 61.656 | 65.410 | 69.957 | 73.166 | 80.0767 |
| 50 | 63,167 | 67.505 | 71.420 | 76.154 | 79-490 | 86.6608 |
| 60 | 74.397 | 79-082 | 83.298 | 88.379 | 91.952 | 99.6072 |
| 70 | 85.527 | 90.531 | 95.023 | 100.425 | 104.215 | 112.3169 |
| 80 | 96.578 | 101.879 | 106.629 | 112.329 | 116.321 | 124.8392 |
| 90 | 107.565 | 113.145 | 118.136 | 124.116 | 128.299 | 137.2083 |
| 100 | 118.498 | 124.342 | 129.561 | 135.807 | 140.169 | 149.4492 |
| 120 | 140.233 | 146.567 | 152.211 | 158.950 | 163.648 | 173.6174 |
| 140 | 161.827 | 168.613 | 174.648 | 181.840 | 186.847 | 197.4508 |
| 200 | 226.021 | 233.994 | 241.058 | 249.445 | 255.207 | 267.5405 |
| 500 | 540.9303 | 553.1268 | 563.8515 | 576.4928 | 585.2066 | 603.4460 |

# جدول القيم الحرجة لتوزيع كاي مربّع

## Chi Square Distribution Table

| d.f. *n* | α | | | | | |
|---|---|---|---|---|---|---|
| | 0.001 | 0.005 | 0.01 | 0.025 | 0.05 | 0.10 |
| 1 | 0.000002 | 0.00004 | 0.00016 | 0.00098 | 0.00393 | 0.0158 |
| 2 | 0.002001 | 0.0100 | 0.0201 | 0.0506 | 0.103 | 0.2107 |
| 3 | 0.02430 | 0.0717 | 0.115 | 0.216 | 0.352 | 0.5844 |
| 4 | 0.09080 | 0.207 | 0.297 | 0.484 | 0.711 | 1.0636 |
| 5 | 0.2102 | 0.412 | 0.554 | 0.831 | 1.145 | 1.6103 |
| 6 | 0.3811 | 0.676 | o.872 | 1.237 | 1.635 | 2.2041 |
| 7 | 0.5985 | 0.989 | 1.239 | 1.690 | 2.167 | 2.8331 |
| 8 | 0.8571 | 1.344 | 1.646 | 2.180 | 2.733 | 3.4895 |
| 9 | 1.152 | 1.735 | 2.088 | 2.700 | 3.325 | 4.1682 |
| 10 | 1.479 | 2.156 | 2.558 | 3.247 | 3.940 | 4.8652 |
| 11 | 1.834 | 2.603 | 3.053 | 3.816 | 4.575 | 5.5778 |
| 12 | 2.214 | 3.074 | 3.571 | 4.404 | 5.226 | 6.3038 |
| 13 | 2.617 | 3.565 | 4.107 | 5.009 | 5.892 | 7.0415 |
| 14 | 3.041 | 4.075 | 4.660 | 5.629 | 6.571 | 7.7895 |
| 15 | 3.483 | 4.601 | 5.229 | 6.262 | 7.261 | 8.5468 |
| 16 | 3.942 | 5.142 | 5.813 | 6.908 | 7.962 | 9.3122 |
| 17 | 4.416 | 5.697 | 6.408 | 7.564 | 8.672 | 10.0852 |
| 18 | 4.905 | 6.265 | 7.015 | 8.231 | 9.390 | 10.8649 |
| 19 | 5.407 | 6.844 | 7.633 | 8.907 | 10.117 | 11.6509 |
| 20 | 5.921 | 7.434 | 8.260 | 9.591 | 10.851 | 12.4426 |
| 21 | 6.447 | 8.034 | 8.897 | 10.283 | 11.591 | 13.2396 |
| 22 | 6.983 | 8.643 | 9.542 | 10.982 | 12.338 | 14.0415 |
| 23 | 7.529 | 9.260 | 10.196 | 11.688 | 13.091 | 14.8480 |
| 24 | 8.085 | 9.886 | 10.856 | 12.401 | 13.848 | 15.6587 |
| 25 | 8.649 | 10.520 | 11.524 | 13.120 | 14.611 | 16.4734 |
| 26 | 9.222 | 11.160 | 12.198 | 13.844 | 15.379 | 17.2919 |
| 27 | 9.803 | 11.808 | 12.878 | 14.573 | 16.151 | 18.1139 |
| 28 | 10.39 | 12.461 | 13.565 | 15,308 | 16.928 | 18.9392 |
| 29 | 10.99 | 13.121 | 14.256 | 16.047 | 17.708 | 19.7677 |
| 30 | 11.59 | 13.787 | 14.953 | 16.791 | 18.493 | 20.5992 |
| 35 | 14.69 | 17.192 | 18.509 | 20.569 | 22.465 | 24.7967 |
| 40 | 17.92 | 20.707 | 22.164 | 24.433 | 26.509 | 29.0505 |
| 45 | 21.25 | 24.311 | 25.901 | 28.366 | 30.612 | 33.3504 |
| 50 | 24.67 | 27.991 | 29.707 | 32.357 | 34.764 | 37.6886 |
| 60 | 31.74 | 35.535 | 37.485 | 40.482 | 43.188 | 46.4589 |
| 70 | 39.04 | 43.275 | 43.275 | 48.758 | 51.739 | 55.3289 |
| 80 | 46.52 | 51.172 | 51.172 | 57.153 | 60.391 | 64.2778 |
| 90 | 54.16 | 59.196 | 59.196 | 65.647 | 69.126 | 73.2911 |
| 100 | 61.92 | 67.328 | 67.328 | 74.222 | 77.929 | 82.3581 |
| 120 | 77.76 | 83.852 | 86.923 | 91.573 | 95.705 | 100.6236 |
| 140 | 93.93 | 100.655 | 104.034 | 109.137 | 113.659 | 119.0293 |
| 200 | 143.84 | 152.241 | 156.432 | 162.728 | 168.279 | 174.8353 |
| 500 | 407.95 | 422.303 | 429.388 | 439.936 | 449.147 | 459.9261 |

## جدول القيم الحرجة لتوزيع فيشر

**Fischer's Distribution Table**

| $n$ | $1-\alpha$ | عدد درجات حرية البسط $m$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 12 | 15 | 20 | 24 | 30 | 60 | 120 | ∞ |
| 1 | 0.90 | 60.2 | 60.7 | 61.2 | 61.7 | 63.0 | 62.3 | 62.8 | 63.1 | 63.3 |
| | 0.95 | 242 | 244 | 246 | 248 | 249 | 250 | 252 | 253 | 254 |
| | 0.975 | 969 | 977 | 985 | 993 | 997 | 1.001 | 1.010 | 1.014 | 1.018 |
| | 0.99 | 6.056 | 6.106 | 6.157 | 6.209 | 6.235 | 6.261 | 6.313 | 6.339 | 6.366 |
| | 0.995 | 24.22 | 24.43 | 24.63 | 24.84 | 24.94 | 25.04 | 25.25 | 25.36 | 25.46 |
| | 0.999 | 605.6 | 610.7 | 615.8 | 620.9 | 623.5 | 626.1 | 631.3 | 636.6 | 634.0 |
| 2 | 0.90 | 9.39 | 9.41 | 9.42 | 9.44 | 9.45 | 9.46 | 9.47 | 9.48 | 9.49 |
| | 0.95 | 19.4 | 19.4 | 19.4 | 19.4 | 19.5 | 19.5 | 19.5 | 19.5 | 19.5 |
| | 0.975 | 39.4 | 39.4 | 39.4 | 39.4 | 39.5 | 39.5 | 39.5 | 39.5 | 39.5 |
| | 0.99 | 99.4 | 99.4 | 99.4 | 99.4 | 99.5 | 99.5 | 99.5 | 99.5 | 99.5 |
| | 0.995 | 199 | 199 | 199 | 199 | 199 | 199 | 199 | 199 | 200 |
| | 0.999 | 999.4 | 999.4 | 999.4 | 999.4 | 999.5 | 999.5 | 999.5 | 999.5 | 999.5 |
| 3 | 0.90 | 5.23 | 5.22 | 5.20 | 5.18 | 5.18 | 5.17 | 5.15 | 5.14 | 5.13 |
| | 0.95 | 8.79 | 8.74 | 8.70 | 8.66 | 8.64 | 8.62 | 8.57 | 8.55 | 8.53 |
| | 0.975 | 14.4 | 14.3 | 14.3 | 14.2 | 14.1 | 14.1 | 14.0 | 13.9 | 13.9 |
| | 0.99 | 27.2 | 27.1 | 26.9 | 26.7 | 26.6 | 26.5 | 26.3 | 26.2 | 26.1 |
| | 0.995 | 43.7 | 43.4 | 43.1 | 42.8 | 42.6 | 42.5 | 42.1 | 42.0 | 41.8 |
| | 0.999 | 129.2 | 128.3 | 127.4 | 126.4 | 125.9 | 125.4 | 124.5 | 124.0 | 123.5 |
| 4 | 0.90 | 3.92 | 3.90 | 3.87 | 3.84 | 3.83 | 3.82 | 3.79 | 3.78 | 3.76 |
| | 0.95 | 5.96 | 5.91 | 5.86 | 5.80 | 5.77 | 5.75 | 5.69 | 5.66 | 5.63 |
| | 0.975 | 8.84 | 8.73 | 8.66 | 8.56 | 8.51 | 8.46 | 8.36 | 8.31 | 8.26 |
| | 0.99 | 14.5 | 14.4 | 14.2 | 14.0 | 13.9 | 13.8 | 13.7 | 13.6 | 13.5 |
| | 0.995 | 21.0 | 20.7 | 20.4 | 20.2 | 20.0 | 19.9 | 19.6 | 19.5 | 19.3 |
| | 0.999 | 48.1 | 47.4 | 46.8 | 46.1 | 45.8 | 45.4 | 44.7 | 44.4 | 44.1 |
| 5 | 0.90 | 3.30 | 3.27 | 3.24 | 3.21 | 3.19 | 3.17 | 3.14 | 3.12 | 3.11 |
| | 0.95 | 4.74 | 4.68 | 4.62 | 4.56 | 4.53 | 4.50 | 4.43 | 4.40 | 4.37 |
| | 0.975 | 6.62 | 6.52 | 6.43 | 6.33 | 6.28 | 6.23 | 6.12 | 6.07 | 6.02 |
| | 0.99 | 10.1 | 9.89 | 9.72 | 9.55 | 9.47 | 9.38 | 9.20 | 9.11 | 9.02 |
| | 0.995 | 13.6 | 13.4 | 13.1 | 12.9 | 12.8 | 12.7 | 12.4 | 12.3 | 12.1 |
| | 0.999 | 26.9 | 26.4 | 25.9 | 25.4 | 25.1 | 24.9 | 24.3 | 24.1 | 23.8 |
| 6 | 0.90 | 2.94 | 2.90 | 2.87 | 2.84 | 2.82 | 2.80 | 2.76 | 2.74 | 2.72 |
| | 0.95 | 4.06 | 4.00 | 3.94 | 3.87 | 3.84 | 3.81 | 3.74 | 3.70 | 3.67 |
| | 0.975 | 5.46 | 5.37 | 5.27 | 5.17 | 5.12 | 5.07 | 4.96 | 4.90 | 4.85 |
| | 0.99 | 7.87 | 7.72 | 7.56 | 7.40 | 7.31 | 7.23 | 7.06 | 6.97 | 6.88 |
| | 0.995 | 10.2 | 10.0 | 9.81 | 9.59 | 9.47 | 9.36 | 9.12 | 9.00 | 8.88 |
| | 0.999 | 18.4 | 18.0 | 17.6 | 17.1 | 16.9 | 16.7 | 16.2 | 16.0 | 15.7 |

# جدول القيم الحرجة لتوزيع فيشر

## Fischer's Distribution Table

| $n$ | $1-\alpha$ | عدد درجات حرية البسط $m$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 0.90 | 39.9 | 49.5 | 53.6 | 55.8 | 57.2 | 58.2 | 58.9 | 59.4 | 59.9 |
| | 0.95 | 161 | 200 | 216 | 225 | 230 | 234 | 237 | 239 | 241 |
| | 0.975 | 648 | 800 | 864 | 900 | 922 | 937 | 948 | 957 | 963 |
| | 0.99 | 4.052 | 5.000 | 5.403 | 5.625 | 5.764 | 5.859 | 5.928 | 5.981 | 6.022 |
| | 0.995 | 16.21 | 20.00 | 21.62 | 22.50 | 23.06 | 23.44 | 23.72 | 23.93 | 24.09 |
| | 0.999 | 405.3 | 500.0 | 540.4 | 562.5 | 576.4 | 585.9 | 592.9 | 598.1 | 602.3 |
| 2 | 0.90 | 8.53 | 9.00 | 9.16 | 9.24 | 9.29 | 9.33 | 9.35 | 9.37 | 9.38 |
| | 0.93 | 18.5 | 19.0 | 19.2 | 19.2 | 19.3 | 19.3 | 19.4 | 19.4 | 19.4 |
| | 0.975 | 38.5 | 39.0 | 39.2 | 39.2 | 39.3 | 39.3 | 39.4 | 39.4 | 39.4 |
| | 0.99 | 98.5 | 99.0 | 99.2 | 99.2 | 99.3 | 99.3 | 99.4 | 99.4 | 99.4 |
| | 0.995 | 199 | 199 | 199 | 199 | 199 | 199 | 199 | 199 | 199 |
| | 0.999 | 998.5 | 999.0 | 999.2 | 999.2 | 999.3 | 999.3 | 999.4 | 999.4 | 999.4 |
| 3 | 0.90 | 5.54 | 5.46 | 5.39 | 5.34 | 3.31 | 5.28 | 5.27 | 5.25 | 5.24 |
| | 0.95 | 10.1 | 9.55 | 9.28 | 9.12 | 9.01 | 8.94 | 8.89 | 8.85 | 8.81 |
| | 0.975 | 17.4 | 16.0 | 15.4 | 15.1 | 14.9 | 14.7 | 14.6 | 14.5 | 14.5 |
| | 0.99 | 34.1 | 30.8 | 29.5 | 28.7 | 28.2 | 27.9 | 27.7 | 27.5 | 27.3 |
| | 0.995 | 55.6 | 49.8 | 47.5 | 46.2 | 45.4 | 44.8 | 44.4 | 44.1 | 43.9 |
| | 0.999 | 167.0 | 148.5 | 141.1 | 137.1 | 134.6 | 132.8 | 131.6 | 130.6 | 129.9 |
| 4 | 0.90 | 4.54 | 4.32 | 4.19 | 4.11 | 4.05 | 4.01 | 3.98 | 3.95 | 3.94 |
| | 0.95 | 7.71 | 6.94 | 6.59 | 6.39 | 6.26 | 6.16 | 6.09 | 6.04 | 6.00 |
| | 0.975 | 12.2 | 10.6 | 9.98 | 9.60 | 9.36 | 9.20 | 9.07 | 8.98 | 8.90 |
| | 0.99 | 21.2 | 18.0 | 16.7 | 16.0 | 15.5 | 15.2 | 15.0 | 14.8 | 14.7 |
| | 0.995 | 31.3 | 26.3 | 24.3 | 23.2 | 22.5 | 22.0 | 21.6 | 21.4 | 21.1 |
| | 0.999 | 74.1 | 61.2 | 56.2 | 53.4 | 51.7 | 50.5 | 49.7 | 49.0 | 48.5 |
| 5 | 0.90 | 4.06 | 3.78 | 3.62 | 3.52 | 3.45 | 3.40 | 3.37 | 3.34 | 3.32 |
| | 0.95 | 6.61 | 5.79 | 5.41 | 5.19 | 5.05 | 4.95 | 4.88 | 4.82 | 4.77 |
| | 0.975 | 10.0 | 8.43 | 7.76 | 7.39 | 7.15 | 6.98 | 6.85 | 6.76 | 6.68 |
| | 0.99 | 16.3 | 13.3 | 12.1 | 11.4 | 11.0 | 10.7 | 10.5 | 10.3 | 10.2 |
| | 0.995 | 22.8 | 18.3 | 16.5 | 15.6 | 14.9 | 14.5 | 14.2 | 14.0 | 13.8 |
| | 0.999 | 47.2 | 37.1 | 33.2 | 31.1 | 29.8 | 28.8 | 28.2 | 27.6 | 27.2 |
| 6 | 0.90 | 3.78 | 3.46 | 3.29 | 3.18 | 3.11 | 3.05 | 3.01 | 2.98 | 2.96 |
| | 0.95 | 5.99 | 5.14 | 4.76 | 4.53 | 4.39 | 4.28 | 4.21 | 4.15 | 4.10 |
| | 0.975 | 8.81 | 7.26 | 6.60 | 6.23 | 5.99 | 5.82 | 5.70 | 5.60 | 5.52 |
| | 0.99 | 13.7 | 10.9 | 9.78 | 9.15 | 8.75 | 8.47 | 8.26 | 8.10 | 7.98 |
| | 0.995 | 18.6 | 14.5 | 12.9 | 12.0 | 11.5 | 11.1 | 10.8 | 10.6 | 10.4 |
| | 0.999 | 35.5 | 27.0 | 23.7 | 21.9 | 20.8 | 20.0 | 19.5 | 19.0 | 18.7 |

## جدول القيم الحرجة لتوزيع فيشر

**Fischer's Distribution Table**

| $n$ | $1-\alpha$ | عدد درجات حرية البسط $m$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 12 | 15 | 20 | 24 | 30 | 60 | 120 | ∞ |
| 7 | 0.90 | 2.70 | 2.67 | 2.63 | 2.59 | 2:58 | 2.56 | 2.51 | 2.49 | 2.47 |
| | 0.95 | 3.64 | 3.57 | 3.51 | 3.44 | 3.41 | 3.38 | 3.30 | 3.27 | 3.23 |
| | 0.975 | 4.76 | 4.67 | 4.57 | 4.47 | 4.42 | 4.36 | 4.25 | 4.20 | 4.14 |
| | 0.99 | 6.62 | 6.47 | 6.31 | 6.16 | 6.07 | 5.99 | 5.82 | 5.74 | 5.65 |
| | 0.995 | 8.38 | 8.18 | 7.97 | 7.75 | 7.65 | 7.53 | 7.31 | 7.19 | 7.08 |
| | 0.999 | 14.1 | 13.7 | 13.3 | 12.9 | 12.7 | 12.5 | 12.1 | 11.9 | 11.7 |
| 8 | 0.90 | 2.54 | 2.50 | 2.46 | 2.42 | 2.40 | 2.38 | 2.34 | 2.32 | 2.29 |
| | 0.95 | 3.35 | 3.28 | 3.22 | 3.15 | 3.12 | 3.08 | 3.01 | 2.97 | 2.93 |
| | 0.975 | 4.30 | 4.20 | 4.10 | 4.00 | 3.95 | 3.89 | 3.78 | 3.73 | 3.67 |
| | 0.99 | 5.81 | 5.67 | 5.52 | 5.36 | 5.28 | 5.20 | 5.03 | 4.95 | 4.86 |
| | 0.995 | 7.21 | 7.01 | 6.81 | 6.61 | 6.50 | 6.40 | 6.18 | 6.06 | 5.95 |
| | 0.999 | 11.5 | 11.2 | 10.8 | 10.5 | 10.3 | 10.1 | 9.73 | 9.53 | 9.33 |
| 9 | 0.90 | 2.42 | 2.38 | 2.34 | 2.30 | 2.28 | 2.25 | 2.21 | 2.18 | 2.16 |
| | 0.95 | 3.14 | 3.07 | 3.01 | 2.94 | 2.90 | 2.86 | 2.79 | 2.75 | 2.71 |
| | 0.975 | 3.96 | 3.87 | 3.77 | 3.67 | 3.61 | 3.56 | 3.43 | 3.39 | 3.33 |
| | 0.99 | 5.26 | 5.11 | 4.96 | 4.81 | 4.73 | 4.65 | 4.48 | 4.40 | 4.31 |
| | 0.995 | 6.42 | 6.23 | 6.03 | 5.83 | 5.73 | 5.62 | 5.41 | 5.30 | 5.19 |
| | 0.999 | 9.89 | 9.57 | 9.24 | 8.90 | 8.72 | 8.55 | 8.19 | 8.00 | 7.81 |
| 10 | 0.90 | 2.32 | 2.28 | 2.24 | 2.20 | 2.18 | 2.16 | 2.11 | 2.08 | 2.06 |
| | 0.95 | 2.98 | 2.91 | 2.84 | 2.77 | 2.74 | 2.70 | 2.62 | 2.58 | 2.54 |
| | 0.975 | 3.72 | 3.62 | 3.52 | 3-42 | 3.37 | 3.31 | 3.20 | 3.14 | 3.08 |
| | 0.99 | 4.85 | 4.71 | 4.56 | 4.41 | 4.33 | 4.25 | 4.08 | 4.00 | 3.91 |
| | 0.995 | 5.85 | 5.66 | 5.47 | 5.27 | 5.17 | 5.07 | 4.86 | 4.75 | 4.64 |
| | 0.999 | 8.75 | 8.45 | 8.13 | 7.80 | 7.64 | 7.47 | 7.12 | 6.94 | 6.76 |
| 12 | 0.90 | 2.19 | 2.15 | 2.10 | 2.06 | 2.04 | 2.01 | 1.96 | 1.93 | 1.90 |
| | 0.95 | 2.75 | 2.69 | 2.62 | 2.54 | 2.51 | 2.47 | 2.38 | 2.34 | 2.30 |
| | 0.975 | 3.37 | 3.28 | 3.18 | 3.07 | 3.02 | 2.96 | 2.85 | 2.79 | 2.72 |
| | 0.99 | 4.30 | 4.16 | 4.01 | 3.86 | 3.78 | 3.70 | 3.54 | 3.45 | 3.36 |
| | 0.995 | 5.09 | 4.91 | 4.72 | 4.53 | 4.43 | 4.33 | 4.12 | 4.01 | 3.90 |
| | 0.999 | 7.29 | 7.00 | 6.71 | 6.40 | 6.25 | 6.09 | 5.76 | 5.59 | 5.42 |
| 15 | 0.90 | 2.06 | 2.02 | 1.97 | 1.92 | 1.90 | 1.87 | 1.82 | 1.79 | 1.76 |
| | 0.95 | 2.54 | 2.48 | 2.40 | 2.33 | 2.29 | 2.25 | 2.16 | 2.11 | 2.07 |
| | 0.975 | 3.06 | 2.96 | 2.86 | 2.76 | 2.70 | 2.64 | 2.52 | 2.46 | 2.40 |
| | 0.99 | 3.80 | 3.67 | 3.52 | 3.37 | 3.29 | 3.21 | 3.05 | 2.96 | 2.87 |
| | 0.995 | 4.42 | 4.25 | 4.07 | 3.88 | 3.79 | 3.69 | 3.48 | 3.37 | 3.26 |
| | 0.999 | 6.08 | 5.81 | 5.54 | 5.25 | 5.10 | 4.95 | 4.64 | 4.48 | 4.31 |

## جدول القيم الحرجة لتوزيع فيشر

## Fischer's Distribution Table

| *n* | $1-\alpha$ | عدد درجات حرية البسط *m* | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 7 | 0.90 | 3.59 | 3.26 | 3.07 | 2.96 | 2.88 | 2.83 | 2.78 | 2.75 | 2.72 |
| | 0.95 | 5.59 | 4.74 | 4.35 | 4.12 | 3.97 | 3.87 | 3.79 | 3.73 | 3.68 |
| | 0.975 | 8.07 | 6.54 | 5.89 | 5.52 | 5.29 | 5.12 | 4.99 | 4.90 | 4.82 |
| | 0.99 | 12.2 | 9.55 | 8.45 | 7.85 | 7.46 | 7.19 | 6.99 | 6.84 | 6.72 |
| | 0.995 | 16.2 | 12.4 | 10.9 | 10.1 | 9.52 | 9.16 | 8.89 | 8.68 | 8.51 |
| | 0.999 | 29.2 | 21.7 | 18.8 | 17.2 | 16.2 | 15.5 | 15.0 | 14.6 | 14.3 |
| 8 | 0.90 | 3.46 | 3.11 | 2.92 | 2.81 | 2.73 | 2.67 | 2.62 | 2.59 | 2.56 |
| | 0.95 | 5.32 | 4.46 | 4.07 | 3.84 | 3.69 | 3.58 | 3.50 | 3.44 | 3.39 |
| | 0.975 | 7.57 | 6.06 | 5.42 | 5.05 | 4.82 | 4.65 | 4.53 | 4.43 | 4.36 |
| | 0.99 | 11.3 | 8.65 | 7.59 | 7.01 | 6.63 | 6.37 | 6.18 | 6.03 | 5.91 |
| | 0.995 | 14.7 | 11.0 | 9.60 | 8.81 | 8.30 | 7.95 | 7.69 | 7.50 | 7.34 |
| | 0.999 | 25.4 | 18.5 | 15.8 | 14.4 | 13.5 | 12.9 | 12.4 | 12.0 | 11.8 |
| 9 | 0.90 | 3.36 | 3.01 | 2.81 | 2.69 | 2.61 | 2.55 | 2.51 | 2.47 | 2.44 |
| | 0.95 | 5.12 | 4.26 | 3.86 | 3.63 | 3.48 | 3.37 | 3.29 | 3.23 | 3.18 |
| | 0.975 | 7.21 | 5.71 | 5.08 | 4.72 | 4.48 | 4.32 | 4.20 | 4.10 | 4.03 |
| | 0.99 | 10.6 | 8.02 | 6.99 | 6.42 | 6.06 | 5.80 | 5.61 | 5.47 | 5.35 |
| | 0.995 | 13.6 | 10.1 | 8.72 | 7.96 | 7.47 | 7.13 | 6.88 | 6.69 | 6.54 |
| | 0.999 | 22.9 | 16.4 | 13.9 | 12.6 | 11.7 | 11.1 | 10.7 | 10.4 | 10.1 |
| 10 | 0.90 | 3.28 | 2.92 | 2.73 | 2.61 | 2.52 | 2.46 | 2.41 | 2.38 | 2.35 |
| | 0.95 | 4.96 | 4.10 | 3.71 | 3.48 | 3.33 | 3.22 | 3.14 | 3.07 | 3.02 |
| | 0.975 | 6.94 | 5.46 | 4.83 | 4.47 | 4.24 | 4.07 | 3.95 | 3.85 | 3.78 |
| | 0.99 | 10.0 | 7.56 | 6.55 | 5.99 | 5.64 | 5.39 | 5.20 | 5.06 | 4.94 |
| | 0.995 | 12.8 | 9.43 | 8.08 | 7.34 | 6.87 | 6.54 | 6.30 | 6.12 | 5.97 |
| | 0.999 | 21.0 | 14.9 | 12.6 | 11.3 | 10.5 | 9.92 | 9.52 | 9.20 | 8.96 |
| 12 | 0.90 | 3.18 | 2.81 | 2.61 | 2.48 | 2.39 | 2.33 | 2.28 | 2.24 | 2.21 |
| | 0.95 | 4.75 | 3.89 | 3.49 | 3.26 | 3.11 | 3.00 | 2.91 | 2.85 | 2.80 |
| | 0.975 | 6.55 | 5.10 | 4.47 | 4.12 | 3.89 | 3.73 | 3.61 | 3.51 | 3.44 |
| | 0.99 | 9.33 | 6.93 | 5.95 | 5.41 | 5.06 | 4.82 | 4.64 | 4.50 | 4.39 |
| | 0.995 | 11.8 | 8.51 | 7.23 | 6.52 | 6.07 | 5.76 | 5.52 | 5.35 | 5.20 |
| | 0.999 | 18.6 | 13.0 | 10.8 | 9.63 | 8.89 | 8.38 | 8.00 | 7.71 | 7.48 |
| 15 | 0.90 | 3.07 | 2.70 | 2.49 | 2.36 | 2.27 | 2.21 | 2.16 | 2.12 | 2.09 |
| | 0.95 | 4.54 | 3.68 | 3.29 | 3.06 | 2.90 | 2.79 | 2.71 | 2.64 | 2.59 |
| | 0.975 | 6.20 | 4.76 | 4.15 | 3.80 | 3.58 | 3.41 | 3.29 | 3.20 | 3.12 |
| | 0.99 | 8.68 | 6.36 | 5.42 | 4.89 | 4.56 | 4.32 | 4.14 | 4.00 | 3.89 |
| | 0.995 | 10.8 | 7.70 | 6.48 | 5.80 | 5.37 | 5.07 | 4.85 | 4.67 | 4.54 |
| | 0.999 | 16.6 | 11.3 | 9.34 | 8.25 | 7.57 | 7.09 | 6.74 | 6.47 | 6.26 |

# جدول القيم الحرجة لتوزيع فيشر

## Fischer's Distribution Table

| $n$ | $1-\alpha$ | عدد درجات حرية البسط $m$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 10 | 12 | 15 | 20 | 24 | 30 | 60 | 120 | ∞ |
| 20 | 0.90 | 1.94 | 1.89 | 1.84 | 1.79 | 1.77 | 1.74 | 1.68 | 1.64 | 1.61 |
| | 0.95 | 2.35 | 2.28 | 2.20 | 2.12 | 2.08 | 2.04 | 1.95 | 1.90 | 1.84 |
| | 0.975 | 2.77 | 2.68 | 2.57 | 2.46 | 2.41 | 2.35 | 2.22 | 2.16 | 2.09 |
| | 0.99 | 3.37 | 3.23 | 3.09 | 2.94 | 2.86 | 2.78 | 2.61 | 2.52 | 2.42 |
| | 0.995 | 3.85 | 3.68 | 3.50 | 3.32 | 3.22 | 3.12 | 2.92 | 2.81 | 2.69 |
| | 0.999 | 5.08 | 4.82 | 4.56 | 4.29 | 4.15 | 4.00 | 3.70 | 3.54 | 3.38 |
| 24 | 0.90 | 1.88 | 1.83 | 1.78 | 1.73 | 1.70 | 1.67 | 1.61 | 1.57 | 1.53 |
| | 0.95 | 2.25 | 2.18 | 2.11 | 2.03 | 1.98 | 1.94 | 1.84 | 1.79 | 1.73 |
| | 0.975 | 2.64 | 2.54 | 2.44 | 2.33 | 2.27 | 2.21 | 2.08 | 2.01 | 1.94 |
| | 0.99 | 3.17 | 3.03 | 2.89 | 2.74 | 2.66 | 2.58 | 2.40 | 2.31 | 2.21 |
| | 0.995 | 3.59 | 3.42 | 3.25 | 3.06 | 2.97 | 2.87 | 2.66 | 2.55 | 2.43 |
| | 0.999 | 4.64 | 4.39 | 4.14 | 3.87 | 3.74 | 3.59 | 3.29 | 3.14 | 2.97 |
| 30 | 0.90 | 1.82 | 1.77 | 1.72 | 1.67 | 1.64 | 1.61 | 1.54 | 1.30 | 1.46 |
| | 0.95 | 2.16 | 2.09 | 2.01 | 1.93 | 1.89 | 1.84 | 1.74 | 1.68 | 1.62 |
| | 0.975 | 2.51 | 2.41 | 2.31 | 2.20 | 2.14 | 2.07 | 1.94 | 1.87 | 1.79 |
| | 0.99 | 2.98 | 2.84 | 2.70 | 2.55 | 2.47 | 2.39 | 2.21 | 2.11 | 2.01 |
| | 0.995 | 3.34 | 3.18 | 3.01 | 2.82 | 2.73 | 2.63 | 2.42 | 2.30 | 2.18 |
| | 0.999 | 4.24 | 4.00 | 3.75 | 3.49 | 3.36 | 3.22 | 2.92 | 2.76 | 2.59 |
| 60 | 0.90 | 1.71 | 1.66 | 1.60 | 1.54 | 1.51 | 1.48 | 1.40 | 1.35 | 1.29 |
| | 0.95 | 1.99 | 1.92 | 1.84 | 1.75 | 1.70 | 1.65 | 1.53 | 1.47 | 1.39 |
| | 0.975 | 2.27 | 2.17 | 2.06 | 1.94 | 1.88 | 1.82 | 1.67 | 1.58 | 1.48 |
| | 0.99 | 2.63 | 2.50 | 2.35 | 2.20 | 2.12 | 2.03 | 1.84 | 1.73 | 1.60 |
| | 0.995 | 2.90 | 2.74 | 2.57 | 2.39 | 2.29 | 2.19 | 1.96 | 1.83 | 1.69 |
| | 0.999 | 3.54 | 3.32 | 3.08 | 2.83 | 2.69 | 2.55 | 2.25 | 2.08 | 1.89 |
| 120 | 0.90 | 1.65 | 1.60 | 1.55 | 1.48 | 1.45 | 1.41 | 1.32 | 1.26 | 1.19 |
| | 0.95 | 1.91 | 1.83 | 1.75 | 1.66 | 1.61 | 1.55 | 1.43 | 1.35 | 1.25 |
| | 0.975 | 2.16 | 2.05 | 1.93 | 1.82 | 1.76 | 1.69 | 1.53 | 1.43 | 1.31 |
| | 0.99 | 2.47 | 2.34 | 2.19 | 2.03 | 1.95 | 1.86 | 1.66 | 1.53 | 1.38 |
| | 0.995 | 2.71 | 2.54 | 2.37 | 2.19 | 2.09 | 1.98 | 1.75 | 1.61 | 1.43 |
| | 0.999 | 3.24 | 3.02 | 2.78 | 2.53 | 2.40 | 1.26 | 1.95 | 1.77 | 1.54 |
| ∞ | 0.90 | 1.60 | 1.55 | 1.49 | 1.42 | 1.38 | 1.34 | 1.24 | 1.17 | 1.00 |
| | 0.95 | 1.83 | 1.75 | 1.67 | 1.57 | 1.52 | 1.46 | 1.32 | 1.22 | 1.00 |
| | 0.975 | 2.05 | 1.94 | 1.83 | 1.71 | 1.64 | 1.57 | 1.39 | 1.27 | 1.00 |
| | 0.99 | 2.32 | 2.18 | 2.04 | 1.88 | 1.79 | 1.70 | 1.47 | 1.32 | 1.00 |
| | 0.995 | 2.52 | 2.36 | 2.19 | 2.00 | 1.90 | 1.79 | 1.53 | 1.36 | 1.00 |
| | 0.999 | 2.96 | 2.74 | 2.51 | 2.27 | 2.13 | 1.99 | 1.66 | 1.45 | 1.00 |

## جدول القيم الحرجة لتوزيع فيشر

## Fischer's Distribution Table

| $n$ | $1-\alpha$ | عدد درجات حرية البسط $m$ | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 20 | 0.90 | 2.97 | 2.59 | 2.38 | 2.25 | 2.16 | 2.09 | 2.04 | 2.00 | 1.95 |
| | 0.95 | 4.35 | 3.49 | 3.10 | 2.87 | 2.71 | 2.60 | 2.51 | 2.45 | 2.39 |
| | 0.975 | 5.87 | 4.46 | 3.86 | 3.51 | 3.29 | 3.13 | 3.01 | 2.91 | 2.84 |
| | 0.99 | 8.10 | 5.85 | 4.94 | 4.43 | 4.10 | 3.87 | 3.70 | 3.56 | 3.46 |
| | 0.995 | 9.94 | 6.99 | 5.82 | 5.17 | 4.76 | 4.47 | 4.26 | 4.09 | 3.96 |
| | 0.999 | 14.8 | 9.95 | 8.10 | 7.10 | 6.46 | 6.02 | 5.69 | 5.44 | 5.24 |
| 24 | 0.90 | 2.93 | 2.54 | 2.33 | 2.19 | 2.10 | 2.04 | 1.98 | 1.94 | 1.91 |
| | 0.95 | 4.26 | 3.40 | 3.01 | 2.78 | 2.62 | 2.51 | 2.42 | 2.36 | 2.30 |
| | 0.975 | 5.72 | 4.32 | 3.72 | 3.38 | 3.15 | 2.99 | 2.87 | 2.78 | 2.70 |
| | 0.99 | 7.82 | 5.61 | 4.72 | 4.22 | 3.90 | 3.67 | 3.50 | 3.36 | 3.26 |
| | 0.995 | 9.55 | 6.66 | 5.52 | 4.89 | 4.49 | 4.20 | 3.99 | 3.83 | 3.69 |
| | 0.999 | 14.0 | 10.6 | 8.52 | 7.39 | 6.68 | 6.18 | 5.82 | 5.54 | 5.31 |
| 30 | 0.90 | 2.88 | 2.49 | 2.28 | 2.14 | 2.05 | 1.98 | 1.93 | 1.88 | 1.85 |
| | 0.95 | 4.17 | 3.32 | 2.92 | 2.69 | 2.53 | 2.42 | 2.33 | 2.27 | 2.21 |
| | 0.975 | 5.57 | 4.18 | 3.59 | 3.25 | 3.03 | 2.87 | 2.75 | 2.65 | 2.57 |
| | 0.99 | 7.56 | 5.39 | 4.51 | 4.02 | 3.70 | 3.47 | 3.30 | 3.17 | 3.07 |
| | 0.995 | 9.18 | 6.35 | 5.24 | 4.62 | 4.23 | 3.95 | 3.74 | 3.58 | 3.45 |
| | 0.999 | 13.3 | 8.77 | 7.05 | 6.12 | 5.53 | 5.12 | 4.82 | 4.58 | 4.39 |
| 60 | 0.90 | 2.79 | 2.39 | 2.18 | 2.04 | 1.95 | 1.87 | 1.82 | 1.77 | 1.74 |
| | 0.95 | 4.00 | 3.15 | 2.76 | 2.53 | 2.37 | 2.25 | 2.17 | 2.10 | 2.04 |
| | 0.975 | 5.29 | 3.93 | 3.34 | 3.01 | 2.79 | 2.63 | 2.51 | 2.41 | 2.33 |
| | 0.99 | 7.08 | 4.98 | 4.13 | 3.65 | 3.34 | 3.12 | 2.95 | 2.82 | 2.72 |
| | 0.995 | 8.49 | 5.80 | 4.73 | 4.14 | 3.76 | 3.49 | 3.29 | 3.13 | 3.01 |
| | 0.999 | 12.0 | 8.65 | 6.81 | 5.82 | 4.76 | 4.37 | 4.09 | 3.87 | 3.69 |
| 120 | 0.90 | 2.75 | 2.35 | 2.13 | 1.99 | 1.90 | 1.82 | 1.77 | 1.72 | 1.68 |
| | 0.95 | 3.92 | 3.07 | 2.68 | 2.45 | 2.29 | 2.18 | 2.09 | 2.02 | 1.96 |
| | 0.975 | 5.15 | 3.80 | 2.23 | 2.89 | 2.67 | 2.52 | 2.39 | 2.30 | 2.22 |
| | 0.99 | 6.85 | 4.79 | 3.95 | 3.48 | 3.17 | 2.96 | 2.79 | 2.66 | 2.56 |
| | 0.995 | 8.18 | 5.54 | 4.50 | 3.92 | 3.55 | 3.28 | 3.09 | 2.93 | 2.81 |
| | 0.999 | 11.4 | 7.32 | 5.79 | 4.95 | 4.42 | 4.04 | 3.77 | 3.55 | 3.38 |
| $\infty$ | 0.90 | 2.71 | 2.30 | 2.08 | 1.94 | 1.85 | 1.77 | 1.72 | 1.67 | 1.63 |
| | 0.95 | 3.84 | 3.00 | 2.60 | 2.37 | 2.21 | 2.10 | 2.01 | 1.94 | 1.88 |
| | 0.975 | 5.02 | 3.69 | 3.12 | 2.79 | 2.57 | 2.41 | 2.29 | 2.19 | 2.11 |
| | 0.99 | 6.63 | 4.61 | 3.78 | 3.32 | 3.02 | 2.80 | 2.64 | 2.51 | 2.41 |
| | 0.995 | 7.88 | 5.30 | 4.28 | 3.72 | 3.35 | 3.09 | 2.90 | 2.74 | 2.62 |
| | 0.999 | 10.8 | 6.91 | 5.42 | 4.62 | 4.10 | 3.74 | 3.47 | 3.27 | 3.30 |

## تابع جدول قيم دالَّة التوزيع الحدَّاني

## Binomial Distribution

$$\sum_{\ell=0}^{k} \mathbf{b}(\ell;n,p) = \sum_{\ell=0}^{k} \binom{n}{\ell} \cdot p^{\ell} \cdot (1-p)^{n-\ell}$$

| n | k | p = 0.10 | 0.20 | 0.25 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 0 | 0.9000 | 0.8000 | 0.7500 | 0.7000 | 0.6000 | 0.5000 | 0.4000 | 0.3000 | 0.2000 | 0.1000 |
| | 1 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 2 | 0 | 0.8100 | 0.6400 | 0.5625 | 0.4900 | 0.3600 | 0.2500 | 0.1600 | 0.0900 | 0.0400 | 0.0100 |
| | 1 | 0.9900 | 0.9600 | 0.9375 | 0.9100 | 0.8400 | 0.7500 | 0.6400 | 0.5100 | 0.3600 | 0.1900 |
| | 2 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 3 | 0 | 0.7290 | 0.5120 | 0.4219 | 0.3430 | 0.2160 | 0.1250 | 0.0640 | 0.0270 | 0.0080 | 0.0010 |
| | 1 | 0.9720 | 0.8960 | 0.8438 | 0.7840 | 0.6480 | 0.5000 | 0.3520 | 0.2160 | 0.1040 | 0.0280 |
| | 2 | 0.9990 | 0.9920 | 0.9844 | 0.9730 | 0.9360 | 0.8750 | 0.7840 | 0.6570 | 0.4880 | 0.2710 |
| | 3 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 4 | 0 | 0.6561 | 0.4096 | 0.3164 | 0.2401 | 0.1296 | 0.0625 | 0.0256 | 0.0081 | 0.0016 | 0.0001 |
| | 1 | 0.9477 | 0.8192 | 0.7383 | 0.6517 | 0.4752 | 0.3125 | 0.1792 | 0.0837 | 0.0272 | 0.0037 |
| | 2 | 0.9963 | 0.9728 | 0.9492 | 0.9163 | 0.8208 | 0.6875 | 0.5248 | 0.3483 | 0.1808 | 0.0523 |
| | 3 | 0.9999 | 0.9984 | 0.9961 | 0.9919 | 0.9744 | 0.9375 | 0.8704 | 0.7599 | 0.5904 | 0.3439 |
| | 4 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 5 | 0 | 0.5905 | 0.3277 | 0.2373 | 0.1681 | 0.0778 | 0.0313 | 0.0102 | 0.0024 | 0.0003 | 0.0000 |
| | 1 | 0.9185 | 0.7373 | 0.6328 | 0.5282 | 0.3370 | 0.1875 | 0.0870 | 0.0308 | 0.0067 | 0.0005 |
| | 2 | 0.9914 | 0.9421 | 0.8965 | 0.8369 | 0.6826 | 0.5000 | 0.3174 | 0.1631 | 0.0579 | 0.0086 |
| | 3 | 0.9995 | 0.9933 | 0.9844 | 0.9692 | 0.9130 | 0.8125 | 0.6630 | 0.4718 | 0.2627 | 0.0815 |
| | 4 | 1.0000 | 0.9997 | 0.9990 | 0.9976 | 0.9898 | 0.9688 | 0.9222 | 0.8319 | 0.6723 | 0.4095 |
| | 5 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 6 | 0 | 0.5314 | 0.2621 | 0.1780 | 0.1176 | 0.0467 | 0.0156 | 0.0041 | 0.0007 | 0.0001 | 0.0000 |
| | 1 | 0.8857 | 0.6554 | 0.5339 | 0.4202 | 0.2333 | 0.1094 | 0.0410 | 0.0109 | 0.0016 | 0.0001 |
| | 2 | 0.9842 | 0.9011 | 0.8306 | 0.7443 | 0.5443 | 0.3438 | 0.1792 | 0.0705 | 0.0170 | 0.0013 |
| | 3 | 0.9987 | 0.9830 | 0.9624 | 0.9295 | 0.8208 | 0.6563 | 0.4557 | 0.2557 | 0.0989 | 0.0159 |
| | 4 | 0.9999 | 0.9984 | 0.9954 | 0.9891 | 0.9590 | 0.8906 | 0.7667 | 0.5798 | 0.3446 | 0.1143 |
| | 5 | 1.0000 | 0.9999 | 0.9998 | 0.9993 | 0.9959 | 0.9844 | 0.9533 | 0.8824 | 0.7379 | 0.4686 |
| | 6 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 7 | 0 | 0.4783 | 0.2097 | 0.1335 | 0.0824 | 0.0280 | 0.0078 | 0.0016 | 0.0002 | 0.0000 | |
| | 1 | 0.8503 | 0.5767 | 0.4449 | 0.3294 | 0.1586 | 0.0625 | 0.0188 | 0.0038 | 0.0004 | 0.0000 |
| | 2 | 0.9743 | 0.8520 | 0.7564 | 0.6471 | 0.4199 | 0.2266 | 0.0963 | 0.0288 | 0.0047 | 0.0002 |
| | 3 | 0.9973 | 0.9667 | 0.9294 | 0.8740 | 0.7102 | 0.5000 | 0.2898 | 0.1260 | 0.0333 | 0.0027 |
| | 4 | 0.9998 | 0.9953 | 0.9871 | 0.9712 | 0.9037 | 0.7734 | 0.5801 | 0.3529 | 0.1480 | 0.0257 |
| | 5 | 1.0000 | 0.9996 | 0.9987 | 0.9962 | 0.9812 | 0.9375 | 0.8414 | 0.6706 | 0.4233 | 0.1497 |
| | 6 | | 1.0000 | 0.9999 | 0.9998 | 0.9984 | 0.9922 | 0.9720 | 0.9176 | 0.7903 | 0.5217 |
| | 7 | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

## تابع جدول قيم دالّة التوزيع الحدّاني

**Binomial Distribution**

$$\sum_{\ell=0}^{k} \mathbf{b}(\ell;n,p)=\sum_{\ell=0}^{k}\binom{n}{\ell}\cdot p^{\ell}\cdot(1-p)^{n-\ell}$$

| $n$ | $k$ | $p$ | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0.10 | 0.20 | 0.25 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 |
| 8 | 0 | 0.4305 | 0.1678 | 0.1001 | 0.0576 | 0.0168 | 0.0039 | 0.0007 | 0.0001 | 0.0000 | |
| | 1 | 0.8131 | 0.5033 | 0.3671 | 0.2553 | 0.1064 | 0.0352 | 0.0085 | 0.0013 | 0.0001 | |
| | 2 | 0.9619 | 0.7969 | 0.6785 | 0.5518 | 0.3154 | 0.1445 | 0.0498 | 0.0113 | 0.0012 | 0.0000 |
| | 3 | 0.9950 | 0.9437 | 0.8862 | 0.8059 | 0.5941 | 0.3633 | 0.1737 | 0.0580 | 0.0104 | 0.0004 |
| | 4 | 0.9996 | 0.9896 | 0.9727 | 0.9420 | 0.8263 | 0.6367 | 0.4059 | 0.1941 | 0.0563 | 0.0050 |
| | 5 | 1.0000 | 0.9988 | 0.9958 | 0.9887 | 0.9502 | 0.8555 | 0.6846 | 0.4482 | 0.2031 | 0.0381 |
| | 6 | | 0.9999 | 0.9996 | 0.9987 | 0.9915 | 0.9648 | 0.8936 | 0.7447 | 0.4967 | 0.1869 |
| | 7 | | 1.0000 | 1.0000 | 0.9999 | 0.9993 | 0.9961 | 0.9832 | 0.9424 | 0.8322 | 0.5695 |
| | 8 | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 9 | 0 | 0.3874 | 0.1342 | 0.0751 | 0.0404 | 0.0101 | 0.0020 | 0.0003 | 0.0000 | | |
| | 1 | 0.7748 | 0.4362 | 0.3003 | 0.1960 | 0.0705 | 0.0195 | 0.0038 | 0.0004 | 0.0000 | |
| | 2 | 0.9470 | 0.7382 | 0.6007 | 0.4628 | 0.2318 | 0.0898 | 0.0250 | 0.0043 | 0.0003 | 0.0000 |
| | 3 | 0.9917 | 0.9144 | 0.8343 | 0.7297 | 0.4826 | 0.2539 | 0.0994 | 0.0253 | 0.0031 | 0.0001 |
| | 4 | 0.9991 | 0.9804 | 0.9511 | 0.9012 | 0.7334 | 0.5000 | 0.2666 | 0.0988 | 0.0196 | 0.0009 |
| | 5 | 0.9999 | 0.9969 | 0.9900 | 0.9747 | 0.9006 | 0.7461 | 0.5174 | 0.2703 | 0.0856 | 0.0083 |
| | 6 | 1.0000 | 0.9997 | 0.9987 | 0.9957 | 0.9750 | 0.9102 | 0.7682 | 0.5372 | 0.2618 | 0.0530 |
| | 7 | | 1.0000 | 0.9999 | 0.9996 | 0.9962 | 0.9805 | 0.9295 | 0.8040 | 0.5638 | 0.2252 |
| | 8 | | | 1.0000 | 1.0000 | 0.9997 | 0.9980 | 0.9899 | 0.9596 | 0.8658 | 0.6126 |
| | 9 | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 10 | 0 | 0.3487 | 0.1074 | 0.0563 | 0.0282 | 0.0060 | 0.0010 | 0.0001 | 0.0000 | | |
| | 1 | 0.7361 | 0.3758 | 0.2440 | 0.1493 | 0.0464 | 0.0107 | 0.0017 | 0.0001 | 0.0000 | |
| | 2 | 0.9298 | 0.6778 | 0.5256 | 0.3828 | 0.1673 | 0.0547 | 0.0123 | 0.0016 | 0.0001 | |
| | 3 | 0.9872 | 0.8791 | 0.7759 | 0.6496 | 0.3823 | 0.1719 | 0.0548 | 0.0106 | 0.0009 | 0.0000 |
| | 4 | 0.9984 | 0.9672 | 0.9219 | 0.8497 | 0.6331 | 0.3770 | 0.1662 | 0.0473 | 0.0064 | 0.0001 |
| | 5 | 0.9999 | 0.9936 | 0.9803 | 0.9527 | 0.8338 | 0.6230 | 0.3669 | 0.1503 | 0.0328 | 0.0016 |
| | 6 | 1.0000 | 0.9991 | 0.9965 | 0.9894 | 0.9452 | 0.8281 | 0.6177 | 0.3504 | 0.1209 | 0.0128 |
| | 7 | | 0.9999 | 0.9996 | 0.9984 | 0.9877 | 0.9453 | 0.8327 | 0.6172 | 0.3222 | 0.0702 |
| | 8 | | 1.0000 | 1.0000 | 0.9999 | 0.9983 | 0.9893 | 0.9536 | 0.8507 | 0.6242 | 0.2639 |
| | 9 | | | | 1.0000 | 0.9999 | 0.9990 | 0.9940 | 0.9718 | 0.8926 | 0.6513 |
| | 10 | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 11 | 0 | 0.3138 | 0.0859 | 0.0422 | 0.0198 | 0.0036 | 0.0005 | 0.0000 | | | |
| | 1 | 0.6974 | 0.3221 | 0.1971 | 0.1130 | 0.0302 | 0.0059 | 0.0007 | 0.0000 | | |
| | 2 | 0.9104 | 0.6174 | 0.4552 | 0.3127 | 0.1189 | 0.0327 | 0.0059 | 0.0006 | 0.0000 | |
| | 3 | 0.9815 | 0.8389 | 0.7133 | 0.5696 | 0.2963 | 0.1133 | 0.0293 | 0.0043 | 0.0002 | |
| | 4 | 0.9972 | 0.9496 | 0.8854 | 0.7897 | 0.5328 | 0.2744 | 0.0994 | 0.0216 | 0.0020 | 0.0000 |
| | 5 | 0.9997 | 0.9883 | 0.9657 | 0.9218 | 0.7535 | 0.5000 | 0.2465 | 0.0782 | 0.0117 | 0.0003 |
| | 6 | 1.0000 | 0.9980 | 0.9924 | 0.9784 | 0.9006 | 0.7256 | 0.4672 | 0.2103 | 0.0504 | 0.0028 |
| | 7 | | 0.9998 | 0.9988 | 0.9957 | 0.9707 | 0.8867 | 0.7037 | 0.4304 | 0.1611 | 0.0185 |
| | 8 | | 1.0000 | 0.9999 | 0.9994 | 0.9941 | 0.9673 | 0.8811 | 0.6873 | 0.3826 | 0.0896 |
| | 9 | | | 1.0000 | 1.0000 | 0.9993 | 0.9941 | 0.9698 | 0.8870 | 0.6779 | 0.3026 |
| | 10 | | | | | 1.0000 | 0.9995 | 0.9964 | 0.9802 | 0.9141 | 0.6862 |
| | 11 | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

## تابع جدول قيم دالَّة التوزيع الحدَّاني

**Binomial Distribution**

$$\sum_{\ell=0}^{k} \mathbf{b}(\ell;n,p)=\sum_{\ell=0}^{k}\binom{n}{\ell}\cdot p^{\ell}\cdot(1-p)^{n-\ell}$$

| n | k | p | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0.10 | 0.20 | 0.25 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 |
| 12 | 0 | 0.2824 | 0.0687 | 0.0317 | 0.0138 | 0.0022 | 0.0002 | 0.0000 | | | |
| | 1 | 0.6590 | 0.2749 | 0.1584 | 0.0850 | 0.0196 | 0.0032 | 0.0003 | 0.0000 | | |
| | 2 | 0.8891 | 0.5583 | 0.3907 | 0.2528 | 0.0834 | 0.0193 | 0.0028 | 0.0002 | 0.0000 | |
| | 3 | 0.9744 | 0.7946 | 0.6488 | 0.4925 | 0.2253 | 0.0730 | 0.0153 | 0.0017 | 0.0001 | |
| | 4 | 0.9957 | 0.9274 | 0.8424 | 0.7237 | 0.4382 | 0.1938 | 0.0573 | 0.0095 | 0.0006 | 0.0000 |
| | 5 | 0.9995 | 0.9806 | 0.9456 | 0.8822 | 0.6652 | 0.3872 | 0.1582 | 0.0386 | 0.0039 | 0.0001 |
| | 6 | 0.9999 | 0.9961 | 0.9857 | 0.9614 | 0.8418 | 0.6128 | 0.3348 | 0.1178 | 0.0194 | 0.0005 |
| | 7 | 1.0000 | 0.9994 | 0.9972 | 0.9905 | 0.9427 | 0.8062 | 0.5618 | 0.2763 | 0.0726 | 0.0043 |
| | 8 | | 0.9999 | 0.9996 | 0.9983 | 0.9847 | 0.9270 | 0.7747 | 0.5075 | 0.2054 | 0.0256 |
| | 9 | | 1.0000 | 1.0000 | 0.9998 | 0.9972 | 0.9807 | 0.9166 | 0.7472 | 0.4417 | 0.1109 |
| | 10 | | | | 1.0000 | 0.9997 | 0.9968 | 0.9804 | 0.9150 | 0.7251 | 0.3410 |
| | 11 | | | | | 1.0000 | 0.9998 | 0.9978 | 0.9862 | 0.9313 | 0.7176 |
| | 12 | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 13 | 0 | 0.2542 | 0.0550 | 0.0238 | 0.0097 | 0.0013 | 0.0001 | 0.0000 | | | |
| | 1 | 0.6213 | 0.2336 | 0.1267 | 0.0637 | 0.0126 | 0.0017 | 0.0001 | 0.0000 | | |
| | 2 | 0.8661 | 0.5017 | 0.3326 | 0.2025 | 0.0579 | 0.0112 | 0.0013 | 0.0001 | | |
| | 3 | 0.9658 | 0.7473 | 0.5843 | 0.4206 | 0.1686 | 0.0461 | 0.0078 | 0.0007 | 0.0000 | |
| | 4 | 0.9935 | 0.9009 | 0.7940 | 0.6543 | 0.3530 | 0.1334 | 0.0321 | 0.0040 | 0.0002 | |
| | 5 | 0.9991 | 0.9700 | 0.9198 | 0.8346 | 0.5744 | 0.2905 | 0.0977 | 0.0182 | 0.0012 | 0.0000 |
| | 6 | 0.9999 | 0.9930 | 0.9757 | 0.9376 | 0.7712 | 0.5000 | 0.2288 | 0.0624 | 0.0070 | 0.0001 |
| | 7 | 1.0000 | 0.9988 | 0.9944 | 0.9818 | 0.9023 | 0.7095 | 0.4256 | 0.1654 | 0.0300 | 0.0009 |
| | 8 | | 0.9998 | 0.9990 | 0.9960 | 0.9679 | 0.8666 | 0.6470 | 0.3457 | 0.0991 | 0.0065 |
| | 9 | | 1.0000 | 0.9999 | 0.9993 | 0.9922 | 0.9539 | 0.8314 | 0.5794 | 0.2527 | 0.0342 |
| | 10 | | | 1.0000 | 0.9999 | 0.9987 | 0.9888 | 0.9421 | 0.7975 | 0.4983 | 0.1339 |
| | 11 | | | | 1.0000 | 0.9999 | 0.9983 | 0.9874 | 0.9363 | 0.7664 | 0.3787 |
| | 12 | | | | | 1.0000 | 0.9999 | 0.9987 | 0.9903 | 0.9450 | 0.7458 |
| | 13 | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 14 | 0 | 0.2288 | 0.0440 | 0.0178 | 0.0068 | 0.0008 | 0.0001 | 0.0000 | | | |
| | 1 | 0.5846 | 0.1979 | 0.1010 | 0.0475 | 0.0081 | 0.0009 | 0.0001 | | | |
| | 2 | 0.8416 | 0.4481 | 0.2811 | 0.1608 | 0.0398 | 0.0065 | 0.0006 | 0.0000 | | |
| | 3 | 0.9559 | 0.6982 | 0.5213 | 0.3552 | 0.1243 | 0.0287 | 0.0039 | 0.0002 | | |
| | 4 | 0.9908 | 0.8702 | 0.7415 | 0.5842 | 0.2793 | 0.0898 | 0.0175 | 0.0017 | 0.0000 | |
| | 5 | 0.9985 | 0.9561 | 0.8883 | 0.7805 | 0.4859 | 0.2120 | 0.0583 | 0.0083 | 0.0004 | |
| | 6 | 0.9998 | 0.9884 | 0.9617 | 0.9067 | 0.6925 | 0.3953 | 0.1501 | 0.0315 | 0.0024 | 0.0000 |
| | 7 | 1.0000 | 0.9976 | 0.9897 | 0.9685 | 0.8499 | 0.6047 | 0.3075 | 0.0933 | 0.0116 | 0.0002 |
| | 8 | | 0.9996 | 0.9978 | 0.9917 | 0.9417 | 0.7880 | 0.5141 | 0.2195 | 0.0439 | 0.0015 |
| | 9 | | 1.0000 | 0.9997 | 0.9983 | 0.9825 | 0.9102 | 0.7207 | 0.4158 | 0.1298 | 0.0092 |
| | 10 | | | 1.0000 | 0.9998 | 0.9961 | 0.9713 | 0.8757 | 0.6448 | 0.3018 | 0.0441 |
| | 11 | | | | 1.0000 | 0.9994 | 0.9935 | 0.9602 | 0.8392 | 0.5519 | 0.1584 |
| | 12 | | | | | 0.9999 | 0.9991 | 0.9919 | 0.9525 | 0.8021 | 0.4154 |
| | 13 | | | | | 1.0000 | 0.9999 | 0.9992 | 0.9932 | 0.9560 | 0.7712 |
| | 14 | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

## تابع جدول قيم دالّة التوزيع الحدّاني

### Binomial Distribution

$$\sum_{\ell=0}^{k} \mathbf{b}(\ell;n,p)=\sum_{\ell=0}^{k}\binom{n}{\ell}\cdot p^{\ell}\cdot(1-p)^{n-\ell}$$

| n | k | p 0.10 | 0.20 | 0.25 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 15 | 0 | 0.2059 | 0.0352 | 0.0134 | 0.0047 | 0.0005 | 0.0000 | | | | |
| | 1 | 0.5490 | 0.1671 | 0.0802 | 0.0353 | 0.0052 | 0.0005 | 0.0000 | | | |
| | 2 | 0.8159 | 0.3980 | 0.2361 | 0.1268 | 0.0271 | 0.0037 | 0.0003 | 0.0000 | | |
| | 3 | 0.9444 | 0.6482 | 0.4613 | 0.2969 | 0.0905 | 0.0176 | 0.0019 | 0.0001 | | |
| | 4 | 0.9873 | 0.8358 | 0.6865 | 0.5155 | 0.2173 | 0.0592 | 0.0093 | 0.0007 | 0.0000 | |
| | 5 | 0.9978 | 0.9389 | 0.8516 | 0.7216 | 0.4032 | 0.1509 | 0.0338 | 0.0037 | 0.0001 | |
| | 6 | 0.9997 | 0.9819 | 0.9434 | 0.8689 | 0.6098 | 0.3036 | 0.0950 | 0.0152 | 0.0008 | |
| | 7 | 1.0000 | 0.9958 | 0.9827 | 0.9500 | 0.7869 | 0.5000 | 0.2131 | 0.0500 | 0.0042 | 0.0000 |
| | 8 | | 0.9992 | 0.9958 | 0.9848 | 0.9050 | 0.6964 | 0.3902 | 0.1311 | 0.0181 | 0.0003 |
| | 9 | | 0.9999 | 0.9992 | 0.9963 | 0.9662 | 0.8491 | 0.5968 | 0.2784 | 0.0611 | 0.0022 |
| | 10 | | 1.0000 | 0.9999 | 0.9993 | 0.9907 | 0.9408 | 0.7827 | 0.4845 | 0.1642 | 0.0127 |
| | 11 | | | 1.0000 | 0.9999 | 0.9981 | 0.9824 | 0.9095 | 0.7031 | 0.3518 | 0.0556 |
| | 12 | | | | 1.0000 | 0.9997 | 0.9963 | 0.9729 | 0.8732 | 0.6020 | 0.1841 |
| | 13 | | | | | 1.0000 | 0.9995 | 0.9948 | 0.9647 | 0.8329 | 0.4510 |
| | 14 | | | | | | 1.0000 | 0.9995 | 0.9953 | 0.9648 | 0.7941 |
| | 15 | | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 16 | 0 | 0.1853 | 0.0281 | 0.0100 | 0.0033 | 0.0003 | 0.0000 | | | | |
| | 1 | 0.5147 | 0.1407 | 0.0635 | 0.0261 | 0.0033 | 0.0003 | 0.0000 | | | |
| | 2 | 0.7892 | 0.3518 | 0.1971 | 0.0994 | 0.0183 | 0.0021 | 0.0001 | | | |
| | 3 | 0.9316 | 0.5981 | 0.4050 | 0.2459 | 0.0651 | 0.0106 | 0.0009 | 0.0000 | | |
| | 4 | 0.9830 | 0.7982 | 0.6302 | 0.4499 | 0.1666 | 0.0384 | 0.0049 | 0.0003 | | |
| | 5 | 0.9967 | 0.9183 | 0.8103 | 0.6598 | 0.3288 | 0.1051 | 0.0191 | 0.0016 | 0.0000 | |
| | 6 | 0.9995 | 0.9733 | 0.9204 | 0.8247 | 0.5272 | 0.2272 | 0.0583 | 0.0071 | 0.0002 | |
| | 7 | 0.9999 | 0.9930 | 0.9729 | 0.9256 | 0.7161 | 0.4018 | 0.1423 | 0.0257 | 0.0015 | 0.0000 |
| | 8 | 1.0000 | 0.9985 | 0.9925 | 0.9743 | 0.8577 | 0.5982 | 0.2839 | 0.0744 | 0.0070 | 0.0001 |
| | 9 | | 0.9998 | 0.9984 | 0.9929 | 0.9417 | 0.7728 | 0.4728 | 0.1753 | 0.0267 | 0.0005 |
| | 10 | | 1.0000 | 0.9997 | 0.9984 | 0.9809 | 0.8949 | 0.6712 | 0.3402 | 0.0817 | 0.0033 |
| | 11 | | | 1.0000 | 0.9997 | 0.9951 | 0.9616 | 0.8334 | 0.5501 | 0.2018 | 0.0170 |
| | 12 | | | | 1.0000 | 0.9991 | 0.9894 | 0.9349 | 0.7541 | 0.4019 | 0.0684 |
| | 13 | | | | | 0.9999 | 0.9979 | 0.9817 | 0.9006 | 0.6482 | 0.2108 |
| | 14 | | | | | 1.0000 | 0.9997 | 0.9967 | 0.9739 | 0.8593 | 0.4853 |
| | 15 | | | | | | 1.0000 | 0.9997 | 0.9967 | 0.9719 | 0.8147 |
| | 16 | | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

## تابع جدول قيم دالّة التوزيع الحدّاني

**Binomial Distribution**

$$\sum_{\ell=0}^{k} \mathbf{b}(\ell;n,p)=\sum_{\ell=0}^{k}\binom{n}{\ell}\cdot p^{\ell}\cdot(1-p)^{n-\ell}$$

| n | k | p: 0.10 | 0.20 | 0.25 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **17** | **0** | 0.1668 | 0.0225 | 0.0075 | 0.0023 | 0.0002 | 0.0000 | | | | |
| | **1** | 0.4818 | 0.1182 | 0.0501 | 0.0193 | 0.0021 | 0.0001 | 0.0000 | | | |
| | **2** | 0.7618 | 0.3096 | 0.1637 | 0.0774 | 0.0123 | 0.0012 | 0.0001 | | | |
| | **3** | 0.9174 | 0.5489 | 0.3530 | 0.2019 | 0.0464 | 0.0064 | 0.0005 | 0.0000 | | |
| | **4** | 0.9779 | 0.7582 | 0.5739 | 0.3887 | 0.1260 | 0.0245 | 0.0025 | 0.0001 | | |
| | **5** | 0.9953 | 0.8943 | 0.7653 | 0.5968 | 0.2639 | 0.0717 | 0.0106 | 0.0007 | 0.0000 | |
| | **6** | 0.9992 | 0.9623 | 0.8929 | 0.7752 | 0.4478 | 0.1662 | 0.0348 | 0.0032 | 0.0001 | |
| | **7** | 0.9999 | 0.9891 | 0.9598 | 0.8954 | 0.6405 | 0.3145 | 0.0919 | 0.0127 | 0.0005 | |
| | **8** | 1.0000 | 0.9974 | 0.9876 | 0.9597 | 0.8011 | 0.5000 | 0.1989 | 0.0403 | 0.0026 | 0.0000 |
| | **9** | | 0.9995 | 0.9969 | 0.9873 | 0.9081 | 0.6855 | 0.3595 | 0.1046 | 0.0109 | 0.0001 |
| | **10** | | 0.9999 | 0.9994 | 0.9968 | 0.9652 | 0.8338 | 0.5522 | 0.2248 | 0.0377 | 0.0008 |
| | **11** | | 1.0000 | 0.9999 | 0.9993 | 0.9894 | 0.9283 | 0.7361 | 0.4032 | 0.1057 | 0.0047 |
| | **12** | | | 1.0000 | 0.9999 | 0.9975 | 0.9755 | 0.8740 | 0.6113 | 0.2418 | 0.0221 |
| | **13** | | | | 1.0000 | 0.9995 | 0.9936 | 0.9536 | 0.7981 | 0.4511 | 0.0826 |
| | **14** | | | | | 0.9999 | 0.9988 | 0.9877 | 0.9226 | 0.6904 | 0.2382 |
| | **15** | | | | | 1.0000 | 0.9999 | 0.9979 | 0.9807 | 0.8818 | 0.5182 |
| | **16** | | | | | | 1.0000 | 0.9998 | 0.9977 | 0.9775 | 0.8332 |
| | **17** | | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| **18** | **0** | 0.1501 | 0.0180 | 0.0056 | 0.0016 | 0.0001 | 0.0000 | | | | |
| | **1** | 0.4503 | 0.0991 | 0.0395 | 0.0142 | 0.0013 | 0.0001 | | | | |
| | **2** | 0.7338 | 0.2713 | 0.1353 | 0.0600 | 0.0082 | 0.0007 | 0.0000 | | | |
| | **3** | 0.9018 | 0.5010 | 0.3057 | 0.1646 | 0.0328 | 0.0038 | 0.0002 | | | |
| | **4** | 0.9718 | 0.7164 | 0.5187 | 0.3327 | 0.0942 | 0.0154 | 0.0013 | 0.0000 | | |
| | **5** | 0.9936 | 0.8671 | 0.7175 | 0.5344 | 0.2088 | 0.0481 | 0.0058 | 0.0003 | | |
| | **6** | 0.9988 | 0.9487 | 0.8610 | 0.7217 | 0.3743 | 0.1189 | 0.0203 | 0.0014 | 0.0000 | |
| | **7** | 0.9998 | 0.9837 | 0.9431 | 0.8593 | 0.5634 | 0.2403 | 0.0576 | 0.0061 | 0.0002 | |
| | **8** | 1.0000 | 0.9957 | 0.9807 | 0.9404 | 0.7368 | 0.4073 | 0.1347 | 0.0210 | 0.0009 | |
| | **9** | | 0.9991 | 0.9946 | 0.9790 | 0.8653 | 0.5927 | 0.2632 | 0.0596 | 0.0043 | 0.0000 |
| | **10** | | 0.9998 | 0.9988 | 0.9939 | 0.9424 | 0.7597 | 0.4366 | 0.1407 | 0.0163 | 0.0002 |
| | **11** | | 1.0000 | 0.9998 | 0.9986 | 0.9797 | 0.8811 | 0.6257 | 0.2783 | 0.0513 | 0.0012 |
| | **12** | | | 1.0000 | 0.9997 | 0.9942 | 0.9519 | 0.7912 | 0.4656 | 0.1329 | 0.0064 |
| | **13** | | | | 1.0000 | 0.9987 | 0.9846 | 0.9058 | 0.6673 | 0.2836 | 0.0282 |
| | **14** | | | | | 0.9998 | 0.9962 | 0.9672 | 0.8354 | 0.4990 | 0.0982 |
| | **15** | | | | | 1.0000 | 0.9993 | 0.9918 | 0.9400 | 0.7287 | 0.2662 |
| | **16** | | | | | | 0.9999 | 0.9987 | 0.9858 | 0.9009 | 0.5497 |
| | **17** | | | | | | 1.0000 | 0.9999 | 0.9984 | 0.9820 | 0.8499 |
| | **18** | | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 |

## تابع جدول قيم دالّة التوزيع الحدّاني

### Binomial Distribution

$$\sum_{\ell=0}^{k} \mathbf{b}(\ell; n, p) = \sum_{\ell=0}^{k} \binom{n}{\ell} \cdot p^{\ell} \cdot (1-p)^{n-\ell}$$

| n | k | p | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 0.10 | 0.20 | 0.25 | 0.30 | 0.40 | 0.50 | 0.60 | 0.70 | 0.80 | 0.90 |
| 19 | 0 | 0.1351 | 0.0144 | 0.0042 | 0.0011 | 0.0001 | | | | | |
| | 1 | 0.4203 | 0.0829 | 0.0310 | 0.0104 | 0.0008 | 0.0000 | | | | |
| | 2 | 0.7054 | 0.2369 | 0.1113 | 0.0462 | 0.0055 | 0.0004 | 0.0000 | | | |
| | 3 | 0.8850 | 0.4551 | 0.2631 | 0.1332 | 0.0230 | 0.0022 | 0.0001 | | | |
| | 4 | 0.9648 | 0.6733 | 0.4654 | 0.2822 | 0.0696 | 0.0096 | 0.0006 | 0.0000 | | |
| | 5 | 0.9914 | 0.8369 | 0.6678 | 0.4739 | 0.1629 | 0.0318 | 0.0031 | 0.0001 | | |
| | 6 | 0.9983 | 0.9324 | 0.8251 | 0.6655 | 0.3081 | 0.0835 | 0.0116 | 0.0006 | | |
| | 7 | 0.9997 | 0.9767 | 0.9225 | 0.8180 | 0.4878 | 0.1796 | 0.0352 | 0.0028 | 0.0000 | |
| | 8 | 1.0000 | 0.9933 | 0.9713 | 0.9161 | 0.6675 | 0.3238 | 0.0885 | 0.0105 | 0.0003 | |
| | 9 | | 0.9984 | 0.9911 | 0.9674 | 0.8139 | 0.5000 | 0.1861 | 0.0326 | 0.0016 | |
| | 10 | | 0.9997 | 0.9977 | 0.9895 | 0.9115 | 0.6762 | 0.3325 | 0.0839 | 0.0067 | 0.0000 |
| | 11 | | 1.0000 | 0.9995 | 0.9972 | 0.9648 | 0.8204 | 0.5122 | 0.1820 | 0.0233 | 0.0003 |
| | 12 | | | 0.9999 | 0.9994 | 0.9884 | 0.9165 | 0.6919 | 0.3345 | 0.0676 | 0.0017 |
| | 13 | | | 1.0000 | 0.9999 | 0.9969 | 0.9682 | 0.8371 | 0.5261 | 0.1631 | 0.0086 |
| | 14 | | | | 1.0000 | 0.9994 | 0.9904 | 0.9304 | 0.7178 | 0.3267 | 0.0352 |
| | 15 | | | | | 0.9999 | 0.9978 | 0.9770 | 0.8668 | 0.5449 | 0.1150 |
| | 16 | | | | | 1.0000 | 0.9996 | 0.9945 | 0.9538 | 0.7631 | 0.2946 |
| | 17 | | | | | | 1.0000 | 0.9992 | 0.9896 | 0.9171 | 0.5797 |
| | 18 | | | | | | | 0.9999 | 0.9989 | 0.9856 | 0.8649 |
| | 19 | | | | | | | 1.0000 | 1.0000 | 1.0000 | 1.0000 |
| 20 | 0 | 0.1216 | 0.0115 | 0.0032 | 0.0008 | 0.0000 | | | | | |
| | 1 | 0.3917 | 0.0692 | 0.0243 | 0.0076 | 0.0005 | 0.0000 | | | | |
| | 2 | 0.6769 | 0.2061 | 0.0913 | 0.0355 | 0.0036 | 0.0002 | | | | |
| | 3 | 0.8670 | 0.4114 | 0.2252 | 0.1071 | 0.0160 | 0.0013 | 0.0000 | | | |
| | 4 | 0.9568 | 0.6296 | 0.4148 | 0.2375 | 0.0510 | 0.0059 | 0.0003 | | | |
| | 5 | 0.9887 | 0.8042 | 0.6172 | 0.4164 | 0.1256 | 0.0207 | 0.0016 | 0.0000 | | |
| | 6 | 0.9976 | 0.9133 | 0.7858 | 0.6080 | 0.2500 | 0.0577 | 0.0065 | 0.0003 | | |
| | 7 | 0.9996 | 0.9679 | 0.8982 | 0.7723 | 0.4159 | 0.1316 | 0.0210 | 0.0013 | 0.0000 | |
| | 8 | 0.9999 | 0.9900 | 0.9591 | 0.8867 | 0.5956 | 0.2517 | 0.0565 | 0.0051 | 0.0001 | |
| | 9 | 1.0000 | 0.9974 | 0.9861 | 0.9520 | 0.7553 | 0.4119 | 0.1275 | 0.0171 | 0.0006 | |
| | 10 | | 0.9994 | 0.9961 | 0.9829 | 0.8725 | 0.5881 | 0.2447 | 0.0480 | 0.0026 | 0.0000 |
| | 11 | | 0.9999 | 0.9991 | 0.9949 | 0.9435 | 0.7483 | 0.4044 | 0.1133 | 0.0100 | 0.0001 |
| | 12 | | 1.0000 | 0.9998 | 0.9987 | 0.9790 | 0.8684 | 0.5841 | 0.2277 | 0.0321 | 0.0004 |
| | 13 | | | 1.0000 | 0.9997 | 0.9935 | 0.9423 | 0.7500 | 0.3920 | 0.0867 | 0.0024 |
| | 14 | | | | 1.0000 | 0.9984 | 0.9793 | 0.8744 | 0.5836 | 0.1958 | 0.0113 |
| | 15 | | | | | 0.9997 | 0.9941 | 0.9490 | 0.7625 | 0.3704 | 0.0432 |
| | 16 | | | | | 1.0000 | 0.9987 | 0.9840 | 0.8929 | 0.5886 | 0.1330 |
| | 17 | | | | | | 0.9998 | 0.9964 | 0.9645 | 0.7939 | 0.3231 |
| | 18 | | | | | | 1.0000 | 0.9995 | 0.9924 | 0.9308 | 0.6083 |
| | 19 | | | | | | | 1.0000 | 0.9992 | 0.9885 | 0.8784 |
| | 20 | | | | | | | | 1.0000 | 1.0000 | 1.0000 |

## جدول قيم العشوائية

**Random Values**

| | |
|---|---|
| 67835 63314 50162 36738<br>25014 15460 11730 76548<br>05393 96770 56087 40875<br>13351 93994 22375 00953<br>85423 60698 01898 61414 | 00551 24909 31894 39333<br>00767 45637 79839 75532<br>28704 04231 40875 13351<br>93994 93994 36009 76548<br>05393 67170 56087 99410 |
| 76548 93547 24769 09404<br>19574 71565 33413 81652<br>45554 27931 05393 96770<br>56087 40875 13351 93994<br>22375 00953 22375 00953 | 33015 19155 11715 48992<br>64998 87080 48660 31288<br>00086 48245 56087 00953<br>22375 00953 99410 16210<br>22375 00953 22375 00953 |
| 00619 85423 60698 01898<br>61414 83525 76548 05393<br>96770 22909 29563 44018<br>64732 93589 61098 04393<br>48245 56087 40875 13351 | 82310 18163 63495 04470<br>78754 90775 10819 56797<br>33751 61414 83525 76548<br>05393 48245 56087 99410<br>04231 40875 13351 93994 |
| 52975 58993 99410 48245<br>56087 41383 31555 12639<br>82667 10324 15457 65515<br>19251 41642 36345 41404<br>81110 64109 09497 76235 | 37774 37953 78837 12538<br>67439 94914 62844 92337<br>99695 99410 48245 56087<br>61414 83525 76548 05393<br>64109 09497 76235 16210 |
| 25640 67257 18671 66658<br>30818 58353 03448 37390<br>96328 04231 40875 13351<br>93994 48245 56087 99410<br>22375 00953 22375 00953 | 50061 42539 14812 07389<br>87891 76255 32072 80083<br>63868 85423 60698 01898<br>58993 99410 48245 56087<br>93994 36009 76548 05393 |
| 60492 78654 51108 66992<br>93183 56920 47196 12452<br>38234 96328 04231 40875<br>13351 41383 31555 12639<br>82667 93994 36009 76548 | 33015 19155 11715 48992<br>64998 87080 48660 31288<br>00086 48245 56087 00953<br>22375 00953 99410 16210<br>22375 00953 22375 00953 |

## جدول قيم العشوائية

**Random Values**

| | |
|---|---|
| 36009 76548 05393 6770<br>56087 40875 13351 93994<br>22375 00953 95892 36962<br>76431 81594 95848 04231<br>13604 75339 01898 61414 | 73152 14511 85285 14536<br>11649 86348 01898 61414<br>83525 76548 05393 96770<br>56087 76431 81594 95848<br>04231 40875 13351 93994 |
| 93547 24769 09404 19574<br>71565 33413 81652 45554<br>27931 12074 98551 37895<br>39941 21225 93629 41469<br>16812 81542 04231 40875 | 12074 98551 37895 39941<br>21225 93629 41469 16812<br>81542 04231 40875 13351<br>93994 36009 76548 05393<br>67701 85423 60698 01898 |
| 41383 31555 12639 82667<br>74624 36348 77319 73408<br>58993 99410 48245 56087<br>22375 00953 22375 00953<br>04231 40875 13351 93994 | 64109 09497 76235 16210<br>89717 65997 02760 24359<br>36009 76548 05393 67705<br>71565 33413 81652 45554<br>56087 40875 13351 93994 |
| 10081 04231 40875 13351<br>93994 73152 14511 85285<br>14536 82454 76810 15941<br>84602 14493 36932 46728<br>71183 16210 64109 09497 | 86576 86944 93296 66456<br>47679 66810 41045 82830<br>47617 76235 16210 16210<br>25014 15460 11730 76548<br>04231 40875 13351 93994 |
| 57550 49620 98480 64451<br>29275 57669 34075 16451<br>42885 58993 99410 48245<br>56087 99410 48245 56087<br>00619 85423 60698 01898 | 68816 37643 19959 46784<br>66125 94932 40350 62533<br>73603 61414 83525 76548<br>05393 04231 40875 13351<br>93994 83525 76548 05393 |
| 19184 64164 66962 40434<br>60602 82175 25471 76107<br>90832 00619 85423 60698<br>01898 48245 56087 61414<br>83525 76548 05393 22375 | 17612 65522 80607 02407<br>06098 92917 55558 15520<br>27038 14536 82454 76810<br>15941 96328 04231 40875<br>13351 76235 16210 16210 |

## جدول كولموغوراف - سميرنوف لاختبار جودة توفيق التوزيعات

**Kolmogorov-Smirnov Table for Test of Goodness of Fit**

| حجم العينة | مستوى الأهمية $\alpha$ | | | | |
|---|---|---|---|---|---|
| $n$ | 0.01 | 0.05 | 0.10 | 0.15 | 0.20 |
| 1 | 0.995 | 0.975 | 0.950 | 0.925 | 0.900 |
| 2 | 0.929 | 0.842 | 0.776 | 0.726 | 0.684 |
| 3 | 0.829 | 0.708 | 0.642 | 0.597 | 0.565 |
| 4 | 0.734 | 0.624 | 0.564 | 0.525 | 0.494 |
| 5 | 0.669 | 0.563 | 0.510 | 0.474 | 0.446 |
| 6 | 0.618 | 0.470 | 0.470 | 0.436 | 0.410 |
| 7 | 0.577 | 0.486 | 0.438 | 0.405 | 0.301 |
| 8 | 0.543 | 0.457 | 0.411 | 0.301 | 0.358 |
| 9 | 0.514 | 0.432 | 0.388 | 0.360 | 0.339 |
| 10 | 0.486 | 0.409 | 0.368 | 0.342 | 0.322 |
| 11 | 0.468 | 0.391 | 0.352 | 0.326 | 0.307 |
| 12 | 0.450 | 0.375 | 0.338 | 0.313 | 0.295 |
| 13 | 0.433 | 0.361 | 0.325 | 0.302 | 0.284 |
| 14 | 0.418 | 0.319 | 0.314 | 0.292 | 0.274 |
| 15 | 0.404 | 0.338 | 0.304 | 0.203 | 0.266 |
| 16 | 0.391 | 0.328 | 0.295 | 0.274 | 0.258 |
| 17 | 0.300 | 0.318 | 0.286 | 0.266 | 0.250 |
| 18 | 0.370 | 0.309 | 0.278 | 0.259 | 0.244 |
| 19 | 0.361 | 0.301 | 0.272 | 0.252 | 0.237 |
| 20 | 0.352 | 0.294 | 0.264 | 0.246 | 0.231 |
| 25 | 0.317 | 0.264 | 0.245 | 0.227 | 0.215 |
| 30 | 0.285 | 0.242 | 0.238 | 0.207 | 0.196 |
| 35 | 0.290 | 0.242 | 0.218 | 0.192 | 0.181 |

لحساب القيمة الحرجة المتعلّقة باختبار كولموغوراف–سميرنوف لجودة التوفيق عندما تكون $35 < n$ تُستخدم العلاقة الآتية:

$$\mathbf{ks}_n = \frac{\sqrt{-0.5 \cdot \mathbf{ln}\left(\frac{\alpha}{2}\right)}}{\sqrt{n}}$$

# ملخص

# لتوزيعات احتمالية شهيرة

**Summary of Famous Probability Distributions**

# توزيعات احتمالية متقطعة

**Discrete Probability Distributions**

| اسم التوزيع | التباين $\mathbf{var}X =$ | الدالة المولدة للعزوم $\mathbf{M}_X(t) =$ | الدالة المميزة $\varphi_X(t) =$ |
|---|---|---|---|
| وحيد النقطة **One point** | $0$ | $\mathbf{e}^{ct}$ | $\mathbf{e}^{c\mathbf{i}t}$ |
| ثنائي النقطة **Tow Point** | $(x_1 - x_2)^2 p(1-p)$ | $p\,\mathbf{e}^{tx_1} + (1-p)\,\mathbf{e}^{tx_2}$ | $p\,\mathbf{e}^{\mathbf{i}tx_1} + (1-p)\,\mathbf{e}^{\mathbf{i}tx_2}$ |
| المنتظم المتقطّع **Discrete Uniform** | $\frac{1}{n}\sum_{k=1}^{n} x_k^2 - \left(\frac{1}{n}\sum_{k=1}^{n} x_k\right)^2$ | $\frac{1}{n}\sum_{k=1}^{n} \mathbf{e}^{tx_k}$ | $\frac{1}{n}\sum_{k=1}^{n} \mathbf{e}^{\mathbf{i}tx_k}$ |
| الحداني **Binomial** | $n\,p(1-p)$ | $\left[1 - p\,(1-\mathbf{e}^{t})\right]^n$ | $\left[1 - p\,(1-\mathbf{e}^{\mathbf{i}t})\right]^n$ |
| الهندسي **Geometric** | $\frac{1-p}{p^2}$ | $\frac{p\,\mathbf{e}^{t}}{1-\mathbf{e}^{t}(1-p)}$ | $\frac{p\,\mathbf{e}^{\mathbf{i}t}}{1-\mathbf{e}^{\mathbf{i}t}(1-p)}$ |
| الحداني السالب **Negative Binomial** | $\frac{n\,p}{(1-p)^2}$ | $\left(\frac{1-p}{1-p\,\mathbf{e}^{t}}\right)^n$ | $\left(\frac{1-p}{1-p\,\mathbf{e}^{\mathbf{i}t}}\right)^n$ |
| فوق الهندسي **Hyper Geometric** | $\frac{nM(N-M)(N-n)}{N^2(N-1)}$ | لا تُعطى بدلالة دوال أولية بسيطة | لا تُعطى بدلالة دوال أولية بسيطة |
| بواسون **Poisson** | $\lambda$ | $\mathbf{e}^{\lambda(\mathbf{e}^{t}-1)}$ | $\mathbf{e}^{\lambda(\mathbf{e}^{\mathbf{i}t}-1)}$ |

# توزيعات احتمالية متقطعة

**Discrete Probability Distributions**

| اسم التوزيع | المعالم<br>Parameters | دالة الكتلة الاحتمالية<br>$P(X=x)=$ | التوقع الرياضياتي<br>$\mathbf{E}X=$ |
|---|---|---|---|
| وحيد النقطة<br>**One point** | $\mathbb{R} \ni c$ | $\mathbf{I}_{\{c\}}(x)$ | $\mathbf{I}_{\{c\}}(x)$ |
| ثنائي النقطة<br>**Tow Point** | $1>p>0$ | $p\cdot\mathbf{I}_{\{x_1\}}(x)+(1-p)\cdot\mathbf{I}_{\{x_2\}}(x)$ | $x_1 p+x_2(1-p)$ |
| المنتظم المتقطّع<br>**Discrete Uniform** | $n\in\mathbb{N}$ | $\frac{1}{n}\cdot\mathbf{I}_{\{x_1,\,x_2,\ldots\ldots,\,x_n\}}(k)$ | $\frac{1}{n}\sum_{i=1}^{n} x_i$ |
| الحدانيّ<br>**Binomial** | $n\in\mathbb{N}$<br>$1>p>0$ | $\binom{n}{x}\cdot p^x\cdot(1-p)^{n-x}\cdot\mathbf{I}_{\mathbb{N}^o_n}(x)$ | $n\cdot p$ |
| الهندسي<br>**Geometric** | $1>p>0$ | $p\cdot(1-p)^{x-1}\cdot\mathbf{I}_{\mathbb{N}}(x)$ | $\frac{1}{p}$ |
| الحدانيّ السالب<br>**Negative Binomial** | $n\in\mathbb{N}$<br>$1>p>0$ | $\binom{x+n-1}{x}\cdot p^n\cdot(1-p)^x\cdot\mathbf{I}_{\mathbb{N}^o}(x)$ | $\frac{n(1-p)}{p}$ |
| فوق الهندسي<br>**Hyper Geometric** | $N,M,n\in\mathbb{N}$<br>$1\le M,n\le N$ | $\binom{M}{x}\cdot\binom{N-M}{n-x}\cdot\binom{N}{n}^{-1}\cdot\mathbf{I}_D(x)$<br>$D:=\mathbb{N}^o_M\cap[\,\max\{0,n-N+M\},\min\{n,M\}\,]$ | $\frac{n\cdot M}{N}$ |
| بواسون<br>**Poisson** | $\lambda>0$ | $\frac{\lambda^x}{x!}\,\mathbf{e}^{-\lambda}\cdot\mathbf{I}_{\mathbb{N}^o}(x)$ | $\lambda$ |

# توزيعات احتمالية مستمرة

**Continuous Probability Distributions**

| اسم التوزيع | التباين $\mathbf{var}X =$ | الدالة المولدة للعزوم $\mathbf{M}_X(t) =$ | الدالة المميزة $\varphi_X(t) =$ |
|---|---|---|---|
| المنتظم المستمر **Continues Uniform** | $\frac{b^2 - a^2}{12}$ | $\frac{\mathbf{e}^{tb} - \mathbf{e}^{ta}}{(b-a)t}$ | $\frac{\mathbf{e}^{\mathbf{i}tb} - \mathbf{e}^{\mathbf{i}ta}}{(b-a)\mathbf{i}t}$ |
| الأسيّ **Exponential** | $\frac{1}{\lambda^2}$ | $\frac{\lambda}{\lambda - t}$ | $\frac{\lambda}{\lambda - \mathbf{i}t}$ |
| الطبيعي **Discrete Uniform** | $\sigma^2$ | $\mathbf{exp}\left[\mu t + \frac{1}{2}t^2\sigma^2\right]$ | $\mathbf{exp}\left[\mu \mathbf{i} t - \frac{1}{2}t^2\sigma^2\right]$ |
| ستودنت **Binomial** | $\frac{n}{n-2}$ $for\ n > 2$ | لا يمكن إعطاؤها بوساطة دوال أوليّة بسيطة | لا يمكن إعطاؤها بوساطة دوال أوليّة بسيطة |
| كايِ مربَّع **Geometric** | $2n$ | $(1+2t)^{-\frac{n}{2}}$ | $(1-2t)^{-\frac{n}{2}}$ |
| فيشر **Negative Binomial** | $\frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$ $for\ n > 4$ | لا يمكن إعطاؤها بوساطة دوال أوليّة بسيطة | لا يمكن إعطاؤها بوساطة دوال أوليّة بسيطة |
| كوشي **Cauchy** | غير موجود | غير موجودة | $\mathbf{exp}\left[\mu \mathbf{i} t - \lambda \lvert t \rvert\right]$ |
| المنطقي **Logistic** | $\frac{\sigma\pi^2}{3}$ | $\mathbf{e}^{\mu t}\ \beta(1-\sigma, 1+\sigma t)$ $for\ \sigma t \in (-1,1)$ | $\mathbf{e}^{\mathbf{i}\mu t}\ \frac{\pi\sigma t}{\mathbf{sinh}(\pi\sigma t)}$ |

# توزيعات احتمالية مستمرة

**Continuous Probability Distributions**

| اسم التوزيع | المعالم<br>Parameters | دالة الكثافة الاحتمالية<br>$f_X(x) =$ | التوقع الرياضياتي<br>$\mathbf{E}X =$ |
|---|---|---|---|
| المنتظم المستمر<br>Continues Uniform | $a, b \in \mathbb{R}$<br>$a < b$ | $\frac{1}{b-a}\,\mathbf{I}_{[a,b]}(x)$ | $\frac{a+b}{2}$ |
| الأسيّ<br>Exponential | $\lambda > 0$ | $\lambda \mathbf{e}^{-\lambda x} \cdot \mathbf{I}_{[0,+\infty)}(x)$ | $\frac{1}{\lambda}$ |
| الطبيعي<br>Discrete Uniform | $\mu \in \mathbb{R}$<br>$\sigma > 0$ | $\frac{1}{\sigma\sqrt{2\pi}}\,\mathbf{exp}\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$ | $\mu$ |
| ستودنت<br>Binomial | $n \in \mathbb{N}$<br>عدد درجات الحرية | $\frac{\Gamma\left(\frac{n+1}{2}\right)\cdot\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n\pi}\cdot\Gamma\left(\frac{n}{2}\right)}$ | $0$ |
| كاي مربَّع<br>Geometric | $n \in \mathbb{N}$<br>عدد درجات الحرية | $\frac{\left(\frac{1}{2}\right)^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}\right)}\,x^{\frac{n}{2}-1}\,\mathbf{e}^{-\frac{x}{2}}\cdot\mathbf{I}_{(0,\infty)}(x)$ | $n$ |
| فيشر<br>Negative Binomial | $n \in \mathbb{N}$<br>$1 > p > 0$ | $\frac{\left(\frac{m}{n}\right)^{\frac{m}{2}} x^{\frac{m}{2}-1}}{\beta\left(\frac{m}{2},\frac{n}{2}\right)}\left(1+\frac{mx}{n}\right)^{-\frac{m+n}{2}}\cdot\mathbf{I}_{[0,\infty)}(x)$ | $\frac{n}{n-2}$<br>$for\ n > 2$ |
| كوشي<br>Cauchy | $\mu \in \mathbb{R}$<br>$\lambda > 0$ | $\frac{\lambda}{\pi\left(\lambda^2+(x-\mu)^2\right)}$ | غير موجود |
| المنطقي<br>Logistic | $\mu \in \mathbb{R}$<br>$\sigma > 0$ | $\frac{\mathbf{exp}\left[-\frac{x-\mu}{\sigma}\right]}{\sigma\left(1+\mathbf{exp}\left[-\frac{x-\mu}{\sigma}\right]\right)^2}$ | $\mu$ |

# المراجع


1-**ALEXANDEROFF, P.S.**; Einfurung in die Mengenlehre und in die allgemeine Topologie; VEB Deutscher-Verlag der Wessenschaften Berlin; **1984**.

2-**ANDERSON & SWEENEY & WILLIAMS**; Statistics for Business and Economics; South Western Educational Publishing; **2011**.

3-**ATHANASIOS PAPOULIS S. & UNNIKRISHNA PILLAI**; Probability, Random Variables And Stochastic Processes Fourth Edition; McGraw-Hill Companies, Inc; **2002**.

4-**ATHREYA, K.B. & LAHIRI, S.N.**; Measure Theory and Probability Theory; Springer Texts in Statistics, Springer-Verlag, NY; **2006**.

5-**BAUER, H.**; Probability Theory and Elements of Measure Theory- English transl; Holt-Rinehart-Winston, NY; **1972**.

6-**BEYER, O. & HACKEL, H. & PIEPER, V. & TIEDGE, J.**; Wahrschichkeitsrechnung und mathematische Statistik; BSB B. G. Teubner Verlagsgesellschaft Leipzig; **1985**.

7-**BHATTACHARYA, R & WAYMIRE, E.C. & AXLER, S. & RIBET, K.A.**; A Basic Course in Probability Theory; Springer Science and Business Media, Inc; **2007**.

8-**BILLINGSLEY, P.**; Probability and Measure, 3rd ed; Wiley, NY; **1995**.

9-**BJÖRN SCHELTER & MATTHIAS WINTERHADLER & JENS TIMMER**; Handbook of Time Series Analysis; WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim; **2006**.

10-**BOROWKOW, A.A.**; Wahrschrinlichkeitstheorie; Alademie Verlag . Berlin; **1976**.

11-**BROCKWELL, P. & DAVIS, R.**; Time Series: Theory and Methods; ; Springer- Verlag New York Inc.; **1987**.

12-**BROCKWELL, P. & DAVIS, R.**; Introduction to Time Series and Forecasting - Second Edition; Springer-Verlag New York, Inc.; **2002**.

13-**CHARLES, W. THERRIEN & MURALI TUMMALA**; Probability for electrical and computer engineers - CRC Press; Roca Raton , New York; **2004**.

14-**CHRISTIAN HESSE**; Angewandte Wahrscheinlichkeitstheorie; Friedr. Vieweg und Sohn Verlagsgesellschaft - Germany; **2003**.

15-**CHRISTOS G. CASSANDRAS & STÉPHANE LAFORTUNE**; Introduction to Discrete Event Systems - Second Edition; Springer Science+Business Media, LLC; **2008**.

16-**DALLMANN, H. & ELSETER, K.-H**; Einfurung in die hohere Mathematik- Band I+II; VEB Gustav Fischer Verlag Jena; **1981**.

17-**DAVID R. ANDERSON & DENNIS J. SWEENEY & THOMAS A. WILLIAMS**; Statistics for Business and Economics - 11e; South-Western, Cengage Learning; **2011**.

18-**De COURSEY, W.J.**; Statistics and Probability for Engineering Applications With Microsoft® Excel; Elsevier Science (USA); **2003**.

19-**DIMITRI P. BERTSEKAS & JOHN N. TSITSIKLIS**; Introduction to Probability; Dimitri P. Bertsekas and John N. Tsitsiklis; **2002**.

20-**DOUGLAS C. MONTGOMERY & GEORGE C. RUNGER**; Applied Statistics and Probability for Engineers -Third Edition; John Wiley & Sons, Inc; **2003**.

21-**DOUGLAS C. MONTGOMERY & CHERYL L. JENNINGS & MURAT KULAHCI**; Introduction to Time Series Analysis and Forecasting; John Wiley & Sons. Inc; **2008**.

22-**EGGERMONT P.P.B. LARICCIA, V.N.**; Maximum Penalized Likelihood Estimation, Volume II: Regression; Springer Science+Business Media, LLC; **2009**.

23-**ENDERTON, H. B.**; Elements of Set Theory; Academic Press College Division New York; **1977**.

24-**FISZ, M.**; Wahrschichkeitsrechnung und mathematische Statistik; VEB Deutscher Verlag der Wissenschaften, Berlin; **1980**.

25-**GEBHARD KIRCHGÄSSER & JÜRGEN WOLTERS**; Introduction to Modern Time Series Analysis; Springer-Verlag Berlin Heidelberg; **2007**.

26-**GNEDENKO, B. W.**;Lehrbuch der Wahrschinlichkeitsrechnung; Akademie – Verlag. Berlin; **1987**.

27-**GNEDENKO,B.W. & LOMONOSSOW & KONIG, D. and another author**; Handbuch der Wahrschinlichkeitsrechnung ( I + II ); Akademie – Verlag. Berlin; **1987**.

28-**GRIMMETT, G.R. & STTRZAKER, D.R**; Probability and Random Processes; Oxford Science Publications; **1982**.

29-**HANS-OTTO GEORGII**; Stochastik; Walter de Gruyter - Berlin . New York; **2002**.

30-**IBRAGIMOV, I.A. & ROZANOV, Y.A.**; Gaussian random processes; Spriger – Verlag – New York-Heidelberg-Berlin; **1970**.

31-**KALLENBERG, O.**; Foundations of Modern Probability, 2nd ed,; Springer-Verlag، NY; **2001**.

32-**KAUSCH, HEIKO**; Zeitreinenprognose; VDM Verlag. München; **2007**.

33-**KRAUSE, B. & METZLER, P.**; Angewadte Statistik; VEB Deutscher-Verlag der Wessenschaften Berlin; **1988**.

34-**KREIβ JENS-PETER**; Einfhrüung in die Zeitreinenanalyse; Verlag Gmbh. Deutschland; **2006**.

35-**KRISTIAN, H.**; Angewandte Wahrscheinlichkeitsthorie; Friedr, Vieweg and Sohn, verlagsgeselschaft mbH, Braunschweig , Wiesbaden; **2003**.

36-**LANDDE ,G .S & SAMBANDHAM ,M.**; Stochastic Versus Deterministic System of differential equation; Marcel Dekker INC. New York – Basel; **2004**.

37-**LARSEN, R. J. & MARX , M. L.**; An introduction to Mathematical Statistics and its Applications; Prentice-Hall International , Inc. USA; **1993**.

**38-LEONID, B. KORALOV & YAKOV, G. SINA**; Theory of Probability and Random Processes - Second Edition; Springer-Verlag Berlin Heidelberg; **2007**.

**39-LOEVE, M.**; Probability Theory ( I + II ); Springer, New York; **1978**.

**40-MAIBAUM, G**; Wahrscheinlichkeitstheorie und mathematische Statistik; VEB Deutscher Verlag der Wessenschaften Berlin; **1980**.

**41-MARVIN, K. SIMON**; Probability Distributions Involving Gaussian Random Variables; Springer Science +Business Media, LLC; **2002**.

**42-MICHEL , H.**; Mass – und Integrationstheorie; VEB Deutscher Verlag der Wessenschaften Berlin; **1978**.

**43-MIKE LEVAN**; Probability and Statistics; McGraw- Hill Companies; **2001**.

**44-RABI BHATTACHARYA & EDWARD C. WAYMIRE**; A Basic Course in Probability Theory; Springer Science+Business Media, Inc.; **2007**.

**45-RASCH, D.**; Einfurung in die mathematische Statistik Band I; VEB Deutscher Verlag der Wissenschaften, Berlin; **1984**.

**46-RENYI, A**; Wahrscheinlichkeitsrechnung; VEB Deutscher Verlag der Wessenschaften Berlin; **1979**.

**47-RICHARD J. LARSEN & MORRIS L. MARX**; Introduction to Mathematical Statistics and Its Applications - 5/E; Pearson • Cloth, 768 pp; **2012**.

**48-ROBERT H. SHUMWAY & DAVID S. STOFFER**; Time Series Analysis and Its Applications; Springer Science+Business Media, LLC; **2006**.

**49-RONALD E. WALPOLE & RAYMOND H. MYERS**; Sharon L. Myers; Keying Ye; Probability & Statistics for Engineers & Scientists – Ninth Edition ; Pearson Education, Inc; **2007**.

**50-ROSSBERG,H.J. & JESIAK,B. & SIEGEL,G.**; Analytic Methods of Probability Theory; Akademie Verlag Berlin; **1985**.

**51-SCNEIDER, I.**; Die Entwicklnug der Wahrscheinlichkeitstheorie von den Anfangen bis **1933**; Akademie-Verlag Berlin; **1989**.

**52-SCOTT L. MILLER & DONALD G. CHILDERS**; Probability and Random Processes; Elsevier Inc; **2004**.

**53-SIRIJAEV,A.N.**; Wahrscheinlichkeit; VEB Deutscher-Verlag der Wessenschaften Berlin; **1988**.

**54-SOONG, T.T.**; Fundamentals of Probability and Statistics for Engineers; John Wiley & Sons Ltd; **2004**.

**55-STORM, R**; Wahrscheinlichkeitsrechnung - mathematische Statistik und statistische Qualitatskontrolle; VEB Fachbuchverlag Leipzig; **1986**.

**56-TOLOSTOW,G.F**; Maβ und Integral; Akademie Verlag Berlin; **1981**.

**57-TREVOR HASTIE & ROBERT TIBSHIRANI & JEROME FRIEDMANN**; The Elements of Statistical Learning - Second Edition; Springer Science+Business Media, LLC; **2009**.

**58-VENKATARAMA KRISHNAN**; Probability and Random Processes; John Wiley & Sons, Inc.; **2006**.

**59-WENTZEL, E.S**; Probability Theory (first steps); Mir Publishers; **1982**.

**60-WINKLER, W**; Vorlesungen zur Mathematischen Statistic; BSB B. G. Teubner Verlagsgesellschaft Leipzig; **1983**.

**61-YRI SUHOV; MARK KELBERT**; Probability and Statistics by Example: II Markov Chains: a Primer in Random Processes and their Applications; Y. Suhov and M. Kelbert; **2008**.

# مراجع عربية

**62- جلال مصطفى الصياد – الاستدلال الإحصائي – منشورات دار المريخ 1993 الرياض.**

**63- حميد عويّد العكله – الجبر الخطي ومبادئ الاحتمالات والإحصاء – لطلاب السّنة الأولى فيزياء – منشورات جامعة البعث 2005 حمص، سوريا.**

**64- حميد عويّد العكله – نظرية الاحتمالات (٢)– لطلاب السّنة الثالثة إحصاء رياضياتي – منشورات جامعة البعث 2007 حمص، سوريا.**

**65- حميد عويّد العكله – نظرية الاحتمالات – لطلاب السّنة الثانية هندسة معلوماتية – منشورات جامعة البعث 2002 حمص، سوريا.**

**66- عبد الله عبد الكريم الشيحة – أسس نظرية التقدير. جامعة الملك سعود – الرياض، المملكة العربية السعودية 1428 هـ.**

**67- فاندل، والتر – السلاسل الزمنية من الوجهة التطبيقية ونماذج بوكس– جنكنز. تعريب عبد المرضي حامد عزام، دار المريخ للنشر، الرياض، المملكة العربية السعودية 1992.**

# معاجم في الرياضيات

**68- علي مصطفى بن الأشتر – معجم الرياضيات – دار النشر أكاديميا – بيروت، لبنان 1995.**

**69- دنان ،ف.م. & باقر، س.ط. & العايدي ، ص.ن. & فران، هـ.ر. & العقيل ، ع.س.هـ. – معجم الرياضيات – منشورات مؤسسة الكويت للتقدّم العلمي – الكويت 1990.**

# دليل الأعلام

## الأسماء رتِّبت بحسب تسلسل ورودها في الكتاب

| | |
|---|---|
| تشبيشيف - رياضياتي روسي | Chebyshev, Pafnuty Lvovich (1821–1894) |
| ثورنيدك - عالم النفس الأمريكي | Thorndike, Edward Lee (1874-1949) |
| بيرسون - إحصائي إنكليزي | Pearson, Karl (1875-1936) |
| بولي - إحصائي واقتصادي بريطاني | Bowley, Sir Arthur Lyon (1869-1957) |
| سبيرمان - إحصائي بريطاني | Spearman, Charles Edward (1863-1945) |
| إنسكومب - إحصائي إنكليزي | Anscombe, Francis John (1918-2001) |
| يول - إحصائي بريطاني | Yule, George Udny (1871 –1951) |
| غالتون - عالم وراثة وإحصائي إنكليزي | Galton, Sir Francis (1822-1911) |
| ليجاندر - رياضياتي فرنسي | Legendre, Adrien-Marie (1752-1833) |
| ويبول - مهندس ورياضياتي سويدي | Weibull, Ernst Hjalmar Waloddi (1887-1979) |
| كلموغوراف - رياضياتي روسي | Kolmogorov, Andrey Nikolaevich (1903-1987) |
| برنشـتاين - رياضياتي روسـي | Bernstein, Sergei Natanovich (1880-1968) |
| بوريل - رياضـي فرنسـي | Borel, Félix Édouard Justin Émile (1871-956) |
| غليفينكو - رياضياتي وفيلسوف أكراني | Glivenko, Valery Ivanovich (1897-1940) |
| شيريايف - رياضياتي وعشوائي روسي | Shiryaev, Albert Nikolaevich (Born 1934) |
| لوفي - إحصائي ورياضياتي فرنسي | Loève, Michel (1907-1979) |
| بوفون - رياضياتي ومؤرخ فرنسي | Buffon, Georges-Louis (1707-1788) |

| | |
|---|---|
| Laplace, Pierre-Simon, marquis (1749-1827) | لابلاس – رياضياتي فرنسي |
| Bertrand, Arthur William Russell (1872-1970) | برتراند – رياضياتي وفيلسوف بريطاني |
| Dirac, Paul Adrien Maurice (1902 –1984) | ديراك – فيزيائي نظري إنكليزي |
| Bayes, Thomas (1701–1761) | بييز – إحصائي وفيلسوف إنكليزي |
| Poisson, Siméon Denis (1781-1840) | بواسون – رياضياتي وفيزيائي فرنسي |
| Radon, Johann Karl August (1887–1956) | رادون – رياضياتي نمساوي |
| Nikodym, Otto Marcin (1887–1974) | نيكوديم – رياضياتي بولوني |
| Bernoulli, Jacob (1655-1705) | برنولي – رياضياتي سويسري |
| Al-Karajī, Abū Bakr ibn Muḥammad ibn al Ḥusayn (953-1029) | الكرخي – فيلسوف ورياضياتي مسلم |
| Newton, Sir Isaac (1642-1726) | نيوتن – رياضياتي وفيزيائي إنكليزي |
| Pascal, Blaise (1623-1662) | باسكال – رياضياتي وفيزيائي فرنسي |
| Moivre, Abraham (1667-1754) | موافير – رياضياتي فرنسي |
| Gauss, Johann Carl Friedrich (1777-1855) | غاوص – رياضياتي ألماني |
| Gosset , William Sealy (1876 –1937) | غوسيت – كيميائي ورياضياتي إنكليزي |
| Cauchy, Augustin-Louis (1789-1857) | كوشي – رياضياتي فرنسي |
| Fisher, Sir Ronald Aylmer (1890 –1962) | فيشر – إحصائي إنكليزي |
| Jensen, Johan Ludwig William (1859-1925) | جينسين – رياضياتي ومهندس دانماركي |
| Schwarz, Karl Hermann Amandus (1843-1921) | شفارتز – فيلسوف ورياضياتي ألماني |
| Markov, Andrey Andreyevich (1856-1922) | ماركوف – رياضياتي روسي |
| Lyapunov, Aleksandr Mikhailovich (1857-1918) | ليابانوف – رياضياتي روسي |
| Bochner, Salomon (1899-1982) | بوخنر – رياضياتي أمريكي (هنغاري الأصل) |
| Lévy, Paul Pierre (1886-1971) | ليفي – رياضياتي فرنسي |
| Gnedenko, Boris Vladimirovich (1912-1995) | غنيدينكو – رياضياتي روسي |
| Tippett, Leonard Henry Caleb (1902-1985) | تيبت – إحصائي إنكليزي |

| | |
|---|---|
| Satterthwaite, F. E. | ساتيرثويت- إحصائي |
| Rao, Calyampudi Radhakrishna (1920) | راو- إحصائي بريطاني (هندي الأصل) |
| Cramér, Harald (1893-1985) | كرامر- رياضياتي إحصائي سويدي |
| Geiger, Johannes Wilhelm (1882-1945) | جايجر- فيزيائي ألماني |
| Marsden, Sir Ernest (1889-1970) | مارسدن- فيزيائي ألماني |
| Smirnov, Nikolai Vasilyevich (1900-1966) | سميرنوف- رياضياتي روسي |
| Anderson, Theodore Wilbur (1918-2016) | أندرسون- رياضياتي وإحصائي أمريكي |
| Darling, Donald Allan (1915-2014) | دارلينغ- إحصائي أمريكي |
| Halberg, Franz (1919-2013) | هالبرغ- رياضياتي فرنسي |
| Carli, Gian Rinaldo (1720-1795) | كارلي- اقتصادي ومؤرخ إيطالي |
| Laspeyres, Ernst Louis Étienne (1834-1913) | لسبيريس- اقتصادي وإحصائي ألماني |
| Paasche, Hermann (1851-1925) | باشي- اقتصادي وإحصائي ألماني |
| Stirling, James (1692-1770) | سترلنغ- رياضي أسكتلندي |
| Fubini, Guido (1879 –1943) | فوبيني- رياضياتي إيطالي |

# ثبت المصطلحات

## English ⇐ عربي

| | | |
|---|---|---|
| | أ | |
| Trend | ٥٦٩ | الاتجاه العام |
| Prior Probability | ٢٢١ | الاحتمال الأسبق |
| Failure Probability | ٣٠٤ | احتمال الفشل |
| Superior Probability | ٢٢١ | الاحتمال اللاحق |
| Success Probability | ٣٠٤ | احتمال النجاح |
| Conditional Probability | ٢١٨ | احتمال شرطي |
| Mathematical Statistics | ١٨٨ | الإحصاء الرياضياتي |
| Statistic | ٤٥١ | إحصاءة |
| Maxwell–Boltzmann Statistic | ١٩٤ | إحصاءة ماكسويل– بولتسمان |
| Sufficiency Statistic | ٤٦٧ | الإحصاءة الكافية |
| Bose – Einstein Statistic | ١٩٤ | إحصاءة بوزي – أينشتاين |
| Fermi-Dirac Statistic | ١٩٤ | إحصاءة فيرمي – ديراك |
| One -Tailed Test | ٥١٨ | اختبار أحادي الطرف (أو أحادي الذيل) |
| Independence Test of two Phenomenon | ٥٤٨ | اختبار استقلال ظاهرتين |
| Optimal Test | ٥٢٠ | الاختبار الأمثل |
| Homogeneity Test of Distribution | ٥٤٤ | اختبار التجانس لتوزيع احتمالي |
| Neumann Deference Test | ٥٨٧ | اختبار الفرق ل نويمان |
| Moore Test–Wallis | ٥٨٨ | اختبار الفروق الأولى ل ووليس- مور |
| Anderson-Darling Test for Goodness of Fit | ٥٤٢ | اختبار أندرسون– دارلينغ لجودة التوفيق |
| Two-Tailed Test | ٥١٨ | اختبار ثنائي الطرف (أو ثنائي الذيل) |
| Goodness of Distribution Fit Test | ٥٣٧ | اختبار جودة توفيق توزيع احتمالي |
| Staurt-Cox Test | ٦٨٨ | اختبار ستراوت-كوكس |
| Non Parametric Test | ٥١٣ | اختبار غير مَعْلمي |
| Statistical Hypothesis test | ٥١٣ | اختبار فرضية إحصائية |

| | | |
|---|---|---|
| Chi-Squared Test for Goodness of Fit | ٣٣٧ | اختبار كاي مربَّع لجودة التوفيق |
| Kolmogorov-Smirnov Test | ٥٣٩ | اختبار كلموغوراف ـ سميرنوف |
| Parametric Test | ٥١٣ | اختبار مَعْلمي |
| Simple Correlation | ١١٧ | الارتباط البسيط |
| Reciprocal Correlation | ١١٧ | الارتباط التبادلي |
| Partial Correlation | ١١٧ | الارتباط الجزئي |
| Simple Linear Correlation | ١١٥ | الارتباط الخطي البسيط |
| Causal Correlation | ١١٧ | الارتباط السببي |
| Multiple Correlation | ١١٧ | الارتباط المتعدّد |
| Spurious Correlation | ١١٧ | الارتباط الوهمي |
| Non Linear Correlation | ١١٨ | الارتباط غير الخطي |
| Statistical Inference | ٤٧٧ | الاستدلال الإحصائي |
| Standardization | ٩٥ | استعيار أو استقياس (أي جعله معيارياً أو قياسياً) |
| Independence of Random Variables | ٣٥١ | استقلال المتغيِّرات العشوائية |
| Independence of Events | ٢٢٣ | استقلال حوادث |
| Absolute Continuation | ٦٦٣ | الاستمرار المطلق |
| Continuous from Below | ٢١٥ | الاستمرار من الأدنى |
| Lower Continuity of a measure | ٦٥٨ | الاستمرار من الأدنى لقياس |
| Continuous from Above | ٢١٥ | الاستمرار من الأعلى |
| Upper Continuity | ٦٥٨ | الاستمرار من الأعلى |
| Framework | ٤٣٩ | الإطار (إطار العمل) |
| Monotony | ٦٥٧ | الاطّراد |
| Five Number and the Box plot | ١٠٧ | الأعداد الخمسة والتمثيل الصندوقي للبيانات |
| Richest Algebra (or Strongest Algebra) | ٢٠٢ | أغنى جبر (أو أقوي جبر) |
| Best Test | ٥١٩ | أفضل اختبار |
| Exponential Regression | ١٦٨ | الانحدار الأسيّ |
| Autoregressive Integrated Moving Average (ARIMA) | ٦٠٤ | الانحدار الذاتي والمتوسطات المتحركة المتكاملة |
| Power Regression | ١٦٩ | انحدار القوى |
| Simple Fractional Regression | ١٦٧ | الانحدار الكسري البسيط |
| Logarithmic Regression | ١٦٧ | الانحدار اللوغاريتمي |
| Mean Deviation (or Average Deviations) | ٨٣ | الانحراف المتوسط (أو متوسط الانحرافات) |
| Standard Deviation | ٨٧ | الانحراف المعياري |
| Relative Standard Deviation | ٩٣ | الانحراف المعياري النسبي |
| Systematic Vibrations | ٥٦٩ | اهتزازات رتيبة |
| Residual | ١٤٥ | الباقي |

| | | |
|---|---|---|
| | ب | |
| Imprint of Time Series | ٥٦٥ | بصمة متسلسلة زمنية |
| Playing Carts | ٢٠٨ | بطاقات لعب |
| Polonium | ٥٣٨ | البولونيوم (مادة مشعّة) |
| Raw Data | ١٢ | بيانات خام |
| Quantitative Data | ٩ | بيانات كميّة |
| Grouped Data | ٢٤ | بيانات مُجمّعة (بيانات مُجدولة) |
| Qualitative Data | ٦ | بيانات نوعيّة |
| | ت | |
| Permutations | ٦٤٥ | تباديل |
| Variance | ٨٥ | التباين |
| Empirical Variance | ٨٦ | التباين العملي (أو تباين العيّنة) |
| Variance of Sample | ٤٥٢ | تباين العيّنة |
| Finite Permutation | ٦٤٥ | تبديل منته |
| Random Experiment | ١٨٨ | تجارب عشوائية |
| Regular Experiment | ١٨٧ | تجارب نظامية |
| Bernoulli Experiment | ٣٠٤ | تجربة برنولية |
| Geiger–Marsden experiment | ٣ | تجربة جايجر ومارسيدن |
| Trivial Partition | ٢٠٣ | التجزئة المبتذلة |
| The Generated Partition by RV | ٣٩٩ | التجزئة المولّدة من متغيّر عشوائي |
| Partition of a Set | ٢٠٣ | تجزئة مجموعة |
| Analysis to Factors | ٤٦٩ | التحليل إلى عوامل |
| Fourier Translation | ٣٩٣ | تحويل فورييه |
| Classification of Samples | ٤٤١ | تصنيف العيّنات |
| Measurable Map | ٦٥٥ | تطبيق قيوس (أو قابل للقياس) |
| Tally | ٣١ | التعداد (تعداد عناصر فئة في جدول تكراري) |
| Classical Definition of Probability | ٢٠٧ | التعريف التقليدي للاحتمال |
| Geometric Definition of Probability | ٢١٠ | التعريف الهندسي للاحتمال |
| Covariance | ١٢٢ | التغاير |
| Total Variation | ١٦٠ | التغيُّر الكلي |
| Explained Variation | ١٦٠ | التغيُّر المُفسَّر |
| Unexplained Variation | ١٦٠ | التغيُّر غير المُفسَّر |
| Seasonal Changes | ٦٠٤ | التغيُّرات الموسميّة |
| Converges with probability equal to one | ٤١٢ | التقارب باحتمال يساوي الواحد |
| Converges in probability | ٤١٢ | التقارب بالاحتمال |
| Converges in Distribution | ٤١٢ | التقارب بالتوزيع |

| | | |
|---|---|---|
| Converges in p-mean | ٤١٢ | التقارب بالمتوسط p |
| Convergence in Square Mean | ٤١٢ | التقارب بالمتوسط التربيعي |
| Converges Almost Sure | ٤١٢ | التقارب تقريباً أكيد (أو شبه أكيد) |
| Point Estimation | ٤٨٥ | التقدير بنقطة (أو التقدير النقطي) |
| Satterthwaite's Approximation | ٤٦٣ | تقريب ستيرثويت |
| Almost Sure | ٢٣٧ | تقريباً أكيد |
| Integral of a Measurable Function | ٦٦١ | تكامل دالّة قيوسة |
| Frequency | ١٣ | تكرار الفئات (عدد العناصر في فئة) |
| Frequency of Class | ٢٤ | تكرار الفئة |
| Ascending Cumulative Frequency | ٣٢ | التكرار المتجمّع الصاعد |
| Descending Cumulative Frequency | ٣٢ | التكرار المتجمّع الهابط |
| Absolut Frequency | ٢٠٥ | التكرار المطلق |
| Percent Frequency | ٣٢ | التكرار المئوي |
| Relative Frequency | ٣١ | التكرار النسبي |
| Expected Frequencies | ٥٤٩ | التكرارات المتوقّعة |
| Representation by Picture | ٢٣ | التمثيل بالصور (الطريقة التصويرية) |
| Pie Chart | ٢١ | تمثيل للبيانات بقرص دائري |
| Dots Representation | ١٦ | تمثيل نقطي |
| Spline | ٢١ | التمهيد الشرائحي |
| Forecasting | ٦١١ | التنبؤ |
| Holt-Winters Smoothing | ٦٠٩ | التنعيم بطريقة هولت-وينتر |
| Smoothing of Time Series | ٦٠١ | تنعيم متسلسلات زمنية |
| Combinations | ٦٤٥ | توافيق |
| Discrete Probability Distribution | ٣٠١ | توزيع احتمالي متقطّع |
| Continuous Probability Distribution | ٣٢٣ | توزيع احتمالي مستمرّ |
| Exponential Distribution | ٣٢٤ | التوزيع الأسي |
| Binomial Distribution | ٢٨٢ | التوزيع الحدّاني |
| Negative Binomial Distribution | ٣١١ | التوزيع الحدّاني السالب |
| Normal (or Gaussian) Distribution | ٣٢٧ | التوزيع الطبيعي (أو التوزيع الغاوسي) |
| Degenerate Distribution | ٣٠١ | التوزيع المضمحل |
| Discrete Uniform Distribution | ٣٠٧ | التوزيع المنتظم المتقطّع |
| Continuous Uniform Distribution | ٣٢٣ | التوزيع المنتظم المستمر |
| Geometric Distribution | ٣١٤ | التوزيع الهندسي |
| Bernoulli Distribution | ٣٠٣ | توزيع برنولي |
| Poisson Distribution | ٢٨٢ | توزيع بواسون |
| Non Symmetric Distributions | ٤٩ | توزيع تكراري غير متناظر |

| | | |
|---|---|---|
| Symmetric Frequency Distribution | ٤٨ | توزيع تكراري متناظر |
| Skewed Frequency Distribution | ٤٩ | توزيع تكراري ملتوي |
| Tow Point Distribution | ٣٠٣ | التوزيع ثنائي النقطة |
| Compound Binomial Distribution | ٢٨٢ | توزيع حداني مُركَّب |
| Negative Skewed Distribution | ٥٠ | توزيع سالب الالتواء |
| Student Distribution (or t-Distribution) | ٣٣٥ | توزيع ستودنت (أو توزيع t) |
| Conditional Distribution | ٢٧٣ | توزيع شرطي |
| Hypergeometric Distribution | ٣١٨ | التوزيع فوق الهندسي (أو الفوهندسي) |
| Fisher distribution | ٣٤٠ | توزيع فيشر |
| Chi Square Distribution | ٣٣٨ | توزيع كاي مربَّع |
| Cauchy Distribution | ٣٣٦ | توزيع كوشي |
| Unimodal Symmetric Distribution | ٤٨ | توزيع متناظر أحادي المنوال |
| Multimodal Symmetric Distribution | ٤٩ | توزيع متناظر متعدّد المنوال |
| Leptokurtic Distribution | ٤٨ | توزيع مدبَّب |
| Compound Distribution | ٢٨١ | توزيع مركَّب |
| Mesokurtic Distribution | ٤٨ | توزيع معتدل |
| Skewed to the Left Distribution | ٥٠ | توزيع ملتوي نحو اليسار |
| Skewed to the Right Distribution | ٤٩ | توزيع ملتوي نحو اليمين |
| Platykurtic Distribution | ٤٨ | توزيع منبسط |
| Positive Skewed Distribution | ٤٩ | توزيع موجب الالتواء |
| Marginal Distribution | ٢٧٨ | توزيع هامشي |
| Weibull Distribution | ٥٤٢ | توزيع وايبول |
| One Point Distribution | ٣٠١ | التوزيع وحيد النقطة |
| Sampling Distributions | ٤٥٦ | توزيعات المعاينة |
| Nets Distribution | ٤٣٤ | توزيعات شبكية |
| Curves Fitting | ١٤٣ | توفيق المنحنيات |
| Fitting a Model of a Time Series | ٦٠٩ | توفيق نموذج متسلسلة زمنية |
| Mathematical Expectation | ٣٥٩ | التوقع الرياضياتي |
| Conditional Expectation | ٣٩٨ | التوقع الشرطي |
| | ج | |
| Algebra | ١٩٥ | جبر |
| Algebra of Events | ١٩٥ | جبر الحوادث |
| Trivial Algebra | ٢٠٢ | الجبر المبتذل |
| Algebra of Finite Complements | ٢٠١ | جبر المتَمّمات المنتهية |
| Pool algebra | ١٩٨ | جبر بول |
| σ–algebra | ١٩٩ | جبر من نوع σ (σ- جبر) |

| | | |
|---|---|---|
| Joint Frequency Table | ١٢٣ | الجدول التكراري المشترك |
| Stem and Leafs Table | ١٤ | جدول الساق والأوراق |
| Frequency Table | ١٣ | جدول تكراري |
| Classes Table | ٢٤ | جدول تكراري ذو فئات |
| Frequency Distribution Table | ٣٠ | جدول توزيع تكراري |
| | ح | |
| Elementary Event | ١٩٥ | حادث ابتدائي |
| Certain Event | ١٩٦ | الحادث الأكيد |
| Impossible Event | ١٩٦ | الحادث المستحيل |
| Simple Event | ١٩٦ | حادثٌ بسيط |
| Atomic Event | ١٩٥ | حادث ذري |
| Almost Certainly Event | ٢١٤ | حادث شبه أكيد |
| Almost Impossible Event | ٢١٠ | حادث شبه مستحيل |
| Size of Sample | ٣ | حجم العيِّنة |
| Size of Population | ٢ | حجم المجتمع |
| Highest Fence (HF) | ٧٧ | الحد الأدنى للقيم المتطرفة بكبرها |
| Lowest Fence (LF) | ٧٧ | الحد الأعلى للقيم المتطرفة بصغرها |
| Lower Confidence Limit | ٤٩٣ | حد الثقة الأدنى |
| Upper Confidence Limit | ٤٩٣ | حد الثقة الأعلى |
| Class Boundaries | ٢٧ | الحدود العمليَّة للفئة |
| Class Limit | ٣٠ | حدود الفئة |
| Borel Field | ٦٥٣ | حقل بوريلي |
| Martingale | ٣٩٨ | الحَكمة |
| Mutually Exclusive Events | ١٩٧ | حوادث متنافية |
| Pairwise Mutually Exclusive Events | ١٩٧ | حوادث متنافية مثنى مثنى |
| Independent Events | ٢٢٣ | حوادث مستقلّة |
| | خ | |
| Consistency Property | ٤٨٨ | خاصيّة الاتساق |
| Regression Line | ٤٠٤ | خط انحدار |
| Curve | ٢١ | خط منحن |
| Error | ١٤٥ | خطأ |
| Standard Error | ١٥٨ | الخطأ المعياري |
| Standard Error of the Estimate | ١٥٨ | الخطأ المعياري للتقدير |
| Error of the First Type | ٥١٥ | الخطأ من النوع الأول |
| Error of the Second Type | ٥١٥ | الخطأ من النوع الثاني |
| Random Walk | ٢٣١ | الخطى العشوائية |

| | د | |
|---|---|---|
| Likelihood Function | ٤٨١ | دالّة الأرجحية |
| Projection Function | ٢٦٦ | دالّة الإسقاط |
| Cumulative Distribution Function | ٢٤٣ | دالّة التوزيع التراكمية |
| Empirical Distribution Function | ٤٥٣ | دالّة التوزيع العمليّة |
| Marginal Distribution Function | ٢٧٩ | دالّة التوزيع الهامشية |
| Probability Mass Function | ٢٥٣ | دالّة الكتلة الاحتمالية |
| Marginal Density Function | ٢٧٨ | دالّة الكثافة الهامشية |
| Radon Nikodym Density Function | ٢٥٨ | دالّة الكثافة ل رادون- نيكوديم |
| Indicator Function | ٦٤٦ | الدالّة المميّزة |
| Indicator Function of an Event | ٢٣٩ | الدالّة المميّزة لحادث |
| Characteristic Function of PD | ٣٩٢ | الدالّة المميّزة للتوزيعات الاحتمالية |
| Characteristic Function of RV | ٣٩٣ | الدالّة المميّزة لمتغيّر عشوائي |
| Indicator Function of a Set | ٦٤٦ | الدالّة المميّزة لمجموعة |
| Logistic Function | ٥٨٤ | الدالّة المنطقية |
| Probability Generating Function | ٣٨٣ | الدالّة المولّدة الاحتمالية |
| Moment Generating Function | ٣٨٩ | الدالّة المولّدة للعزوم |
| Growth Function | ٥٨٣ | دالّة النمو |
| Simple Function | ٦٤٧ | دالّة بسيطة |
| Beta Function | ٦٤٩ | دالة بيتا |
| Step Function | ٦٤٧ | دالّة درجيّة |
| Sample Function | ٤٥١ | دالّة عيّنة |
| Gama Function | ٦٤٨ | دالة غاما |
| Measurable Function | ٦٥٥ | دالّة قيوسة (أو قابلة للقياس) |
| Convex Function | ٦٤٧ | دالّة محدّبة |
| Concave Function | ٦٤٧ | دالّة مقعّرة |
| Degree | ٢٢ | درجة (تستخدم لقياس الزوايا) |
| $\tau$-Score | ٩٥ | الدرجة المعيارية -$\tau$ (الدرجة المعيارية التائية) |
| $z$-Score | ٩٤ | الدرجة المعيارية -$z$ |
| Standard Score | ٩٣ | درجة معيارية |
| Accuracy of Estimate | ١٥٨ | دقة التقدير |
| Base Period | ٦١٥ | دورة الأساس |
| Current Period | ٦١٦ | الدورة الحالية (أو الدورة الجارية) |
| Compare Period | ٦٠٦ | دورة المقارنة |
| Seasonal Cycle | ٥٧٣ | الدورة الموسميّة |
| Previous Cycle | ٦١١ | دورة سابقة |

| | | |
|---|---|---|
| Irregular Cyclical | ٥٧٠ | دورية غير منتظمة (عشوائية) |
| | ذ | |
| Atom of the Partition | ٢٠٣ | ذرة تجزئة |
| | ر | |
| Anscombe's Quartet | ١٣٢ | رباعية إنسكومب |
| Quartile (Quartiles) | ٧٣ | ربيعي (ربيعيات) |
| First Quartile (or Lower Quartile $Q_1$) | ٧٣ | الربيعي الأول (أو الربيعي الأدنى) |
| Third Quartile (or Upper Quartile $Q_3$) | ٧٣ | الربيعي الثالث (أو الربيعي الأعلى) |
| Second Quartile | ٧٣ | الربيعي الثاني |
| Rank | ٧٤ | رتبة |
| Order of the Partial Correlation | ١٧٨ | رتبة الارتباط الجزئي |
| Index Number | ٦١٨ | رقم قياسي |
| Simple Index Number | ٦١٨ | رقم قياسي بسيط |
| Simple Aggregate Index Number | ٦٢٣ | رقم قياسي كلي بسيط |
| Weighted Index Numbers | ٦٢٥ | رقم قياسي موزون |
| Composite Index Number | ٦٢٣ | رقماً قياسياً مركباً |
| | س | |
| Pulling without Replacement | ١٩٣ | السحب بدون إرجاع |
| Draw without Replacement | ٤٤٠ | السحب بدون إرجاع (أو بدون إعادة) |
| Draw with Replacement | ٤٤٠ | السحب مع الإرجاع (أو مع الإعادة) |
| Pulling with Replacement | ١٩٢ | السحب مع الإرجاع |
| Angular Frequency | ٥٩٢ | السرعة الزاوية للاهتزاز |
| Default Class width | ٢٨ | السعة الافتراضية (لفئة في جدول تكراري) |
| Length of Class | ٢٤ | سعة الفئة (في جدول تكراري) |
| | ش | |
| Markov Condition | ٤٨٨ | شرط ماركوف |
| Tail | ١٩٠ | شعار (في تجربة قذف قطعة نقود معدنية) |
| Whiskers | ١٠٨ | الشُعيْرتان (شُعَيرتا التمثيل الصندوقي للبيانات) |
| | ص | |
| Monotone class | ٦٥٤ | صف مطّرد |
| Head | ٤ | صورة (يقابلها الكتابة على قطع النقود المعدنية) |
| Total Probability Formula | ٢٢٠ | صيغة الاحتمال التام |
| Steiner Formula | ٣٧٤ | صيغة شتاينر |
| | ض | |
| White Noise | ٥٧٠ | الضجيج الأبيض |

| | | |
|---|---|---|
| | ط | |
| Methods of Sampling | ٤٤٠ | طرائق سحب العيِّنة |
| Maximum Likelihood Method | ٤٧٨ | طريقة الأرجحية العظمى |
| Halberg's Cosine Method | ٥٩١ | طريقة التجيب ل هالبرغ |
| Tables Method | ١٣ | طريقة الجداول |
| Moments Method | ٤٧٨ | طريقة العزوم |
| Successive Differences Method | ٥٧٨ | طريقة الفروق المتتالية |
| Least Square Method | ١٤٤ | طريقة المربعات الصغرى |
| Ratio to Moving-Average Method | ٦٠٥ | طريقة النسبة إلى متوسط متحرك |
| Winters' Method | ٦٠٠ | طريقة وينتر |
| Winter's Method for Generate Forecasts | ٦٠٨ | طريقة وينتر لتوليد التنبوَات |
| White Noise Process | ٥٧٠ | طوري ضجيج أبيض |
| Random Process | ٥٦٤ | طوري عشوائي |
| Real Random Process | ٥٦٤ | طوري عشوائي حقيقي |
| Point Process | ٢٣٨ | طوري نقطي |
| Length of Period | ٥٩٢ | طول الدور |
| | ع | |
| Default Number | ٢٨ | العدد الافتراضي (لفئات جدول تكراري) |
| Number of Classes | ٣٠ | عدد الفئات |
| Box plot | ١٠٧ | العرض (أو التمثيل) الصندوقي |
| Graphing of Data | ١١ | عرض البيانات |
| Bar Graph | ١٧ | العرض الشرائطي |
| The $k^{th}$ Factorial Moment | ٣٧٦ | العزم العاملي من المرتبة k |
| Centered Moment | ٣٧٢ | العزم المركزي |
| Central Moment of Order r | ٩٦ | العزم المركزي من المرتبة r |
| Factorial Moment | ٣٧٥ | العزوم العامليّ (أو الضربيّ) |
| Moments of random variable | ٣٥٨ | العزوم لمتغيِّر عشوائي |
| Central Moment | ٩٦ | عزوم مركزي |
| Causal Chain Relation | ١١٧ | علاقة السلسلة السببية |
| Reciprocal Relation | ١١٧ | علاقة تبادلية |
| Unidirectional Relation | ١١٦ | علاقة وحيد الاتجاه |
| Spurious Relation | ١١٧ | علاقة وهمية |
| Statistics | ١ | علم الإحصاء |
| Stochastic | ١٨٨ | علم العشوائيات |
| Initial Element | ٤٤٥ | عنصر البدء |
| Random Element | ٢٣٨ | عنصر عشوائي |

| | | |
|---|---|---|
| Intentional Samples | ٤٤١ | عيّنات عمديّة (أو قصديّة) |
| Selected Sample | ٤٤٧ | عيّنات منتقاة |
| Sample | ٢ | عيّنة |
| Systematic Sample | ٤٤٥ | العيّنة الرتيبة (أو المنتظمة) |
| Simple Random Sample | ٤٤٤ | العيّنة العشوائية البسيطة |
| Quota Sample | ٤٧٧ | عيّنة حصص (مبنية على المحاصصة) |
| Stratified Sample | ٤٤٥ | عيّنة طبقية |
| Random Sample | ٤٤١ | عيّنة عشوائية |
| Cluster Sample | ٤٤٦ | عيّنة عنقودية |
| High-Size Sample | ٤٤٧ | عيّنة كبيرة الحجم |
| | غ | |
| Uncountable Set | ٦٤٢ | غير قابلة للعدّ |
| Uncorrelated | ٣٦٨ | غير مرتبطين |
| | ف | |
| Confidence Interval | ٤٩٢ | فترة ثقة |
| Lag | ٥٦٤ | فجوة زمنية |
| Statistical hypothesis | ٥١٣ | فرضية إحصائية |
| Null Hypothesis | ٥١٤ | الفرضية الابتدائية (أو الفرضية الصفرية) |
| Alternative Hypothesis | ٥١٤ | الفرضية البديلة |
| Simple hypothesis | ٥١٨ | فرضية بسيطة |
| Composite hypothesis | ٥١٨ | فرضية مركّبة |
| Acrophase | ٥٩٢ | فرق الصفحة |
| Separable | ٢٣٨ | فَصول (أو قابل للفصل) |
| Probability Space | ١٨٧ | فضاء احتمالي |
| Space of Elementary Events | ١٩٥ | فضاء الحوادث الابتدائية |
| Measure Space | ٦٥٦ | فضاء قياس |
| Metric Space | ٦٥٣ | فضاء متري |
| Measurable Space | ٢١٤ | فضاء مقيس |
| Class (Group, Category or Interval) | ٢٤ | الفئة (في الجداول التكرارية) |
| | ق | |
| Countable Set | ٦٤٢ | قابلة للعدّ |
| Integrablity of a Real Function | ٦٦١ | قابلية المكاملة لدالّة حقيقية |
| Additive Rule | ٦٤٦ | قاعدة الجمع |
| Multiplication Rule | ٦٤٦ | قاعدة الضرب |
| Chebyshev Empirical Rule | ٨٩ | قاعدة تشبيشيف التجريبيّة |
| Addition Law | ٢١٥ | قانون الجمع |

| | | |
|---|---|---|
| Multiplication Law of Probability | ٢٢٠ | قانون الضرب في الاحتمالات |
| Weak Law of Large Numbers | ٤١٤ | القانون الضعيف للأعداد الكبيرة |
| Strong Law of Large Numbers | ٤١٤ | القانون القوي للأعداد الكبيرة |
| Bernoulli's Law of Large Numbers | ٤١٦ | قانون برنولي للأعداد الكبيرة |
| Poisson's Law of Large Numbers | ٤١٦ | قانون بواسون للأعداد الكبيرة |
| Distribution Law | ٢٤٠ | قانون توزيع |
| Cardinality of a Set | ٦٤٢ | قدرة مجموعة |
| Coin | ١٩٠ | قطعة نقود معدنية |
| Laws of Large Numbers | ٤١٤ | قوانين الأعداد الكبيرة |
| Power of the Test | ٥١٥ | قوة الاختبار (أو حساسية الاختبار) |
| Ogive | ٤٣ | القوس القوطي |
| Measure | ٦٥٦ | قياس |
| σ-Finite Measure | ٦٥٧ | قياس σ- منتهٍ |
| Probability Measure | ١٨٩ | قياس احتمالي |
| Measuring Fitting Accuracy | ٦١١ | قياس جودة التوفيق |
| Dirac Measure | ٢١٥ | قياس ديراك |
| Lebesgue Measure | ٦٥٧ | قياس لوبيغ |
| Finite Measure | ٦٥٧ | قياس منته |
| Observation Values | ٤٤٩ | قيم ملحوظة (أو مشاهدة) |
| Critical Value | ٥١٦ | قيمة حرجة |
| Extreme Value | ٧٧ | قيمة متطرفة |
| Predicted Value | ١٤٦ | قيمة مقدَّرة |
| Outlier Value | ٧٧ | قيمة منعزلة |
| Fitted Value | ١٤٦ | قيمة موفَّقة (تمَّ توفيقها من منحنى التوفيق) |
| Relative Value | ٦١٨ | قيمة نسبية |
| | ك | |
| Probability Mass | ٢٥٣ | كتلة احتمالية |
| Relative Efficiency | ٤٨٩ | الكفاءة النسبية |
| Integrable | ٢٥٩ | كمول (أو قابل للمكاملة) |
| Fisher's Information Quantity | ٤٧١ | كمية معلومات فيشر |
| | ل | |
| Scatter Plot | ١١٩ | لوحة الانتشار |
| | م | |
| Pauli Principle | ١٩٤ | مبدأ باولي (مبدأ الاستبعاد) |
| Fisher -Neumann Proposition | ٤٦٨ | مبرهنة فيشر – نويمان |
| Chebyshev Inequality | ٣٨١ | متباينة تشبيشيف |

| | | |
|---|---|---|
| Jensen inequality | ٣٦٣ | متباينة جينسين |
| Rao-Cramer Inequality | ٤٩٠ | متباينة راو- كرامر |
| Cauchy – Schwartz inequality | ٣٦٤ | متباينة كوشي– شفارتز |
| Cauchy–Schwartz Inequality | ٣٦٤ | متباينة كوشي– شفارتز |
| Lyapunov Inequality | ٣٧١ | متباينة ليابانوف |
| Markov inequality | ٣٧١ | متباينة ماركوف |
| Essentially Sequence | ٤١٣ | متتالية أساسية |
| Cauchy Sequence | ٤١٣ | متتالية كوشي |
| Random Vector | ٢٣٨ | متّجه عشوائي |
| Discrete Random Vector | ٢٧٢ | متّجه عشوائي متقطّع |
| Mixed Random Vector | ٢٧٤ | متّجه عشوائي مختلط |
| Continuous Random Vectors | ٢٧٣ | متّجه عشوائي مستمر |
| Time Series | ٥٦٣ | متسلسلة زمنية |
| Variable | ٥ | متغيّر |
| Response Variable | ١١٥ | متغيّر الاستجابة |
| Dependent Variable | ١١٥ | المتغيّر التابع |
| Predictor Variable | ١٤٦ | المتغيّر المُتنبأ به |
| Explanatory Variable | ١١٥ | متغيّر تفسيري |
| Random Variable | ٢٣٣ | متغيّر عشوائي |
| Bernoulli Random Variable | ٣٠٤ | متغيّر عشوائي برنولي |
| Complex Random Variable | ٢٣٨ | متغيّر عشوائي عقدي |
| Discrete Random Variable | ٢٥٢ | متغيّر عشوائي متقطّع |
| Continuous Random Variable | ٢٥٨ | متغيّر عشوائي مستمر |
| Normalized random variable | ٣٧٥ | متغيّر عشوائي مستنظم |
| Regular Random variable | ٤٧١ | متغيّر عشوائي نظامي |
| Quantitative Variable | ٧ | متغيّر كميّ |
| Discrete Variable | ٦ | متغيّر متقطّع |
| Continuous Variable | ٦ | متغيّر مستمرّ |
| Confusing Variable | ١٧٨ | متغيّر مشوّش |
| Perturbation Variable | ١٧٨ | متغيّر مشوّش |
| Independent Random Variables | ٣٥١ | متغيّرات عشوائية مستقلّة |
| Independent Variables | ٣٥١ | متغيّرات مستقلّة |
| Qualitative Variables | ٧ | متغيّرات نوعيّة |
| Discrete | ٤ | متقطّع (أو منفصل) |
| Complement of an Event | ١٩٦ | متمّم حادث |
| Mean (or Arithmetic Mean) | ٥٦ | المتوسط (أو المتوسط الحسابي) |

| | | |
|---|---|---|
| Harmonic Mean | ٦٤ | المتوسط التوافقي |
| Trimmed mean (TrMean) | ٧٢ | المتوسط الحسابي المشذّب |
| Mean of Sample | ٤٥٢ | متوسط العيّنة |
| Mean Absolute Deviation (MAD) | ٦١١ | متوسط القيم المطلقة للانحرافات |
| Mean Absolute Percentage Error (MAPE) | ٦١١ | متوسط القيم المطلقة للنسب المئوية للأخطاء |
| Weighted Mean | ٦١ | المتوسط الموزون |
| Geometric Mean | ٦٢ | المتوسط الهندسي |
| Weighted Geometric Mean | ٦٣ | المتوسط الهندسي الموزون |
| Moving Average | ٦٠٢ | متوسط متحرك |
| Mean Square Deviation (MSD) | ٦١١ | متوسط مربّع الانحرافات |
| Mean Squared Error | ١٦٠ | متوسط مربّع الخطأ |
| The Risk of Type II | ٥١٥ | المجازفة من النوع الثاني |
| Population | ١ | مجتمع إحصائي |
| Independent Populations | ٤٦١ | مجتمعات مستقلّة |
| Borel Set | ٢١٤ | مجموعة بوريلية |
| Compact Set | ٢٣٩ | مجموعة متراصة |
| Convex Set | ٦٤٤ | مجموعة محدّبة |
| Continuous Set | ٤ | مجموعة مستمرّة (أو متصلة) |
| Random Closed Set | ٢٣٩ | مجموعة مغلقة عشوائية |
| Set of Outcomes | ١٨٩ | مجموعة نتائج |
| Bertrand Paradox | ٢١١ | مُحيِّرة برتراند |
| Leptokurtic | ٤٨ | مُدبَّب |
| Relative Frequency Histogram | ٣٧ | مدرَّج التكراري النسبي |
| Frequency Histogram | ٣٦ | مدرَّج تكراري |
| Percentile Frequency Histogram | ٣٨ | مدرَّج تكراري مئوي |
| Range | ٢٧ | المدى |
| Interquartile Range (IQR) | ٩١ | المدى الربيعي |
| Trend Component | ٥٦٩ | مُركَّبة الاتجاه العام |
| Random Error Component | ١٥٥ | مركبة الخطأ العشوائي |
| Period (or Cyclical) Component | ٥٧٠ | المُركّبة الدورية |
| Seasonal Component | ٥٦٩ | المُركّبة الموسميّة |
| Systematic Component | ١٥٥ | المُركّبة النظامية (أو الرتيبة) |
| Irregular Component | ٥٧٠ | مُركّبة عشوائية (أو غير منتظمة) |
| Midpoint | ٣١ | مركز الفئة (في جدول تكراري) |
| Gnedenko Problem | ٤٣٤ | مسألة غنيدينكو |
| Lyapunov Problem | ٤٢٨ | مسألة ليابانوف |

| | | |
|---|---|---|
| Lindenberg – Feller Problem | ٤٣١ | مسألة ليندنبرغ – فيلر |
| Lindenberg – Levy Problem | ٤٢١ | مسألة ليندنبرغ – ليفي |
| Moivre – Laplace Problem | ٤١٧ | مسألة موافير – لابلاس |
| Independent | ٢٢٣ | مستقل |
| Statistical Independent | ٢٢٤ | مستقلّة إحصائياً |
| Stochastic Independent | ٢٢٤ | مستقلّة عشوائياً |
| Pairwise Independent | ٢٢٤ | مستقلّة مثنى مثنى |
| Straight Regression | ١٤٥ | مستقيم الانحدار |
| Absolutely Continuous | ٢٥٨ | مستمر مطلقاً |
| Level | ٦٠٩ | المستوى (أو التمهيد) |
| Confidence Level | ٤٩٣ | مستوى الثقة (أو معامل الثقة) |
| Significance Level | ٥١٥ | مستوى أهمية |
| Significance Level of the Test | ٥١٥ | مستوى أهمية الاختبار |
| Comprehensive Survey | ٢ | المسح الشامل |
| Covariance Matrix | ١٢٥ | مصفوفة التغاير |
| Polygon | ١٩ | مضلّع (تمثيل بقطع مستقيمة متتالية) |
| Ascending Cumulative Frequency Paragon (Less than Ogive) | ٤٢ | مضلّع التكرار المتجمّع الصاعد (أقل من قوسلقوطي) |
| Descending Cumulative Frequency Polygon (Less than Ogive) | ٤٢ | مضلّع التكرار المتجمّع الهابط (أقل من قوس قوطي) |
| Frequency Polygon | ٤١ | مضلّع تكراري |
| Cumulative Frequency Paragon | ٤٢ | مضلّعات تكراري تراكمي |
| Amplitude | ٥٩٢ | مطال |
| Regression Equation | ٤٠٤ | معادلة انحدار |
| Welch–Satterthwaite Equation | ٤٦٣ | معادلة ويلش ستيرثويت |
| Partial Correlation Coefficient | ١٧٨ | مُعامل الارتباط الجزئي |
| Coefficient of Association | ١٤٠ | مُعامل الاقتران |
| Moment Coefficient of Skewness | ١٠٠ | مُعامل الالتواء العزومي |
| Coefficient of Determination | ١٣٩ | مُعامل التحديد |
| Coefficient of Dispersion | ٩٣ | مُعامل التشتت |
| Coefficient of Variation | ٩٢ | مُعامل التغيُّر |
| Coefficient of Contingency | ١٤١ | مُعامل التوافق |
| Roughness Coefficient | ٦٠١ | مُعامل الخشونة |
| Bowley's Coefficient of Skewness | ١٠٢ | مُعامل بولي للالتواء |
| Person's Coefficient of Correlation | ١٢٧ | مُعامل بيرسون للارتباط |
| Person's Coefficient of Skewness | ١٠٠ | مُعامل بيرسون للالتواء |
| Person's Coefficient of Kurtosis | ١٠٦ | مُعامل بيرسون للتفلطح |
| Spearman Rho Correlation Coefficient | ١٣١ | مُعامل سبيرمان رو للارتباط |

| | | |
|---|---|---|
| Spearman's Rank Correlation Coefficient | ١٣٠ | مُعامل سبيرمان لارتباط الرتب |
| Regression Coefficients | ١٤٦ | مُعاملات (أو وسطاء) الانحدار |
| Correlation Coefficients | ١٢١ | مُعاملات الارتباط |
| Simple Average | ٦١٩ | المعدَّل البسيط |
| Parameter | ٥٦ | مَعْلَمَة (أو وسيط) |
| Trend Parameter | ٦١٠ | مَعْلَمَة الاتجاه العام |
| Seasonal Parameter | ٦١٠ | مَعْلَمَة الموسمي |
| Population Parameter | ٤٧٧ | مَعْلَمَة مجتمع إحصائي |
| Perturbation Parameter | ٥١٨ | مَعْلَمَة مشوَّشة |
| Asymptotically Normal Distributed | ٤١٩ | مقارب للتوزيع الطبيعي |
| Resistance | ٤ | مقاومة أوميَّة |
| Descriptive Measures | ٥٥ | مقاييس وصفية |
| Unbiasedness Quantity | ٤٨٦ | مقدار التحيُّز |
| Statistical Estimator | ٤٧٨ | مقدِّر إحصائي |
| Maximum Likelihood Estimator | ٤٨١ | مُقدِّر الأرجحية العظمى |
| Highest Efficient Estimator | ٤٨٩ | المقدِّر الأعلى كفاءة |
| Estimator more efficient from another | ٤٨٩ | المقدِّر الأكفأ من مُقدِّر آخر |
| Pooled Estimator of Variance | ٤٦٣ | المقدِّر المشترك للتباين |
| Unbiased Estimator | ٤٨٦ | المقدِّر المُنصِف (أو غير المنحاز) |
| Efficient Estimator | ٤٨٩ | مقدِّر كفوء |
| Consistent Estimator | ٤٨٨ | مقدِّر متَّسق |
| Asymptotic Unbiased Estimator | ١٨٧ | مقدِّر منصف تقاربيًا |
| Nominal Scale | ٧ | مقياس اسمي |
| Variability Measure | ٨٣ | مقياس الاختلاف (تسمية أخرى لمقياس التشتت) |
| Spread Measures | ٨٣ | مقياس الانتشار(تسمية أخرى لمقياس التشتت) |
| Scatter Measure | ٨٣ | مقياس التبعثر(تسمية أخرى لمقياس التشتت) |
| Skewness Measure | ٩٩ | مقياس التواء |
| Form Measures | ٩٦ | مقياس الشكل |
| Interval Scale | ٩ | مقياس الفترة |
| Ration Scale | ٩ | مقياس النسبة |
| Ordinal Scale | ٨ | مقياس ترتيبي |
| Dispersion Measure | ٨٢ | مقياس تشتت |
| Kurtosis Measure | ١٠٥ | مقياس تفلطح (أو تفرطح) |
| Position Measure | ٥٥ | مقياس موضع |
| Central Tendency Measure | ٥٦ | مقياس نزعة مركزيَّة |
| Observation | ٣ | ملحوظة (أو مشاهدة) |

| | | |
|---|---|---|
| Frequency Curve | ٤٣ | منحنى تكراري |
| Unbiasedness | ٤٨٦ | منصف أو غير متحيز |
| Critical Region of the Test | ٥١٥ | المنطقة الحرجة للاختبار |
| Rejection Area | ٥١٥ | منطقة الرفض |
| Optimum Rejection Area | ٥١٩ | منطقة الرفض المثلى |
| Acceptance Area | ٥١٥ | منطقة القبول |
| Critical Region | ٥١٥ | منطقة حرجة |
| Mode | ٧٠ | المنوال |
| Positive Definite | ٣٩٤ | موجبة التعيين |
| Fisher's ideal index | ٦٣١ | المؤشّر المثالي ل فيشر |
| Paasche index | ٦٢٥ | مؤشّر باشّي |
| Laspeyres Index | ٦٢٥ | مؤشّر لسبيريس |
| Percentile (Percentiles) | ٧٧ | منين (المنينات) |
| r-Percentile (or $r^{th}$ Percentile) | ٧٧ | المنين الرائي |
| | ن | |
| Outcome | ١٨٩ | نتيجة (نتيجة تجربة) |
| Taylor's Expansion | ٣٩٠ | نشر تايلور |
| Stirling Expansion | ٦٤٥ | نشر سترلنغ |
| Subadditive | ٦٥٧ | نصف جمعي |
| $\sigma$–Finite | ٦٥٧ | نصف جمعي σ- منتهٍ (أو نصف جمعي عدود) |
| Finite Subadditive | ٦٥٧ | نصف جمعي منتهٍ |
| Continuity theorem's | ٣٩٢ | نظريات الاستمرار |
| Convolution theorem's | ٣٩٢ | نظريات التلاف |
| Inversion theorems | ٣٩٢ | نظريات العكس |
| Central limit theorem's | ٤١١ | نظريات النهاية المركزية |
| Uniqueness theorem's | ٣٩٢ | نظريات الوحدانية |
| Decision-Making Theory | ٥١٣ | نظرية اتخاذ القرار |
| Probability Theory | ١٨٨ | نظرية الاحتمالات |
| Estimation Theory | ٤٧٧ | نظرية التقدير |
| Martingale Theorem | ٣٩٨ | نظرية الحَكَمة |
| Queuing theory | ٣٢٠ | نظرية الطوابير |
| Measure Theory | ١٥٩ | نظرية القياس |
| Sampling Theory | ٣ | نظرية المعاينة |
| Local Central Limit Theorem | ٤٣٤ | نظرية النهاية المركزية الموضعية |
| Reliability Theory | ٣٢٠ | نظرية الوثوقية |
| Bochner Theorem | ٣٩٦ | نظرية بوخنر |

| | | |
|---|---|---|
| Beppo – Levi's Theorem | ٦٦٢ | نظرية بيبو – ليفي |
| Bays's theorem | ٢٢١ | نظرية بييز |
| Radon–Nikodym Theorem | ٦٣٣ | نظرية رادون – نيكوديم |
| Gnedenko Theorem | ٤٣٤ | نظرية غنيدينكو |
| Fubini Theorem | ٦٦٢ | نظرية فوبيني |
| Lyapunov Theorem | ٤٢٨ | نظرية ليابانوف |
| Lévy Theorem | ٣٩٧ | نظرية ليفي |
| Lindenberg – Feller Theorem | ٤٣١ | نظرية ليندنبرغ – فيلر |
| Lindenberg–Levy Theorem | ٤٢٢ | نظرية ليندنبرغ– ليفي |
| Neyman Pearson Theorem | ٥١٩ | نظرية نيمان – بيرسون |
| Critical Point | ٥١٦ | نقطة حرجة |
| Harmonic Model | ٥٧٠ | نموذج (أو نمط) توافقي |
| Additive Model | ٥٧٢ | نموذج الجمع |
| Multiplicative Model | ٥٧٢ | نموذج الضرب |
| Smoothing Model | ٥٧٢ | نموذج تنعيم |
| Theoretical Mathematical Model | ٤٥٠ | نموذج رياضياتي نظري |
| Random Kern | ٢٣٩ | نواة عشوائية |
| | هـ | |
| Stochastic Geometry | ٢١٣ | الهندسة العشوائية |
| | و | |
| Fahrenheit | ١٠ | وحدة الفهرنهايت لقياس الحرارة |
| Kelvin | ١٠ | وحدة الكلفين لقياس الحرارة المنخفضة جداً |
| Sampling Unit | ٤٣٩ | وِحْدَة المعاينة |
| Celsius | ١٠ | وحدة لقياس الحرارة (الدرجة المئوية) |
| Monetary Units | ٦١٩ | وحدة نقدية |
| Weight for the Trend | ٦٠٩ | الوزن من أجل الاتجاه العام |
| Weight for the seasonal component | ٦٠٩ | الوزن من أجل المركّبة الموسميّة |
| Weight for the level | ٦٠٩ | الوزن من أجل المستوى (أو من أجل التمهيد) |
| Median | ٦٦ | وسيط (أو الوسط) |
| | ي | |
| Equal Almost Every Where | ٦٦٢ | يساوي تقريباً في كل مكان |

# English ⇒ عربي


| | | |
|---|---|---|
| | **A** | |
| Absolut Frequency | ٢٠٥ | التكرار المطلق |
| Absolute Continuation | ٦٦٣ | الاستمرار المطلق |
| Absolutely Continuous | ٢٥٨ | مستمر مطلقاً |
| Acceptance Area | ٥١٥ | منطقة القبول |
| Accuracy of Estimate | ١٥٨ | دقة التقدير |
| Acrophase | ٥٩٢ | فرق الصفحة |
| Addition Law | ٢١٥ | قانون الجمع |
| Additive Model | ٥٧٢ | نموذج الجمع |
| Additive Rule | ٦٤٦ | قاعدة الجمع |
| Algebra | ١٩٥ | جبر |
| Algebra of Events | ١٩٥ | جبر الحوادث |
| Algebra of Finite Complements | ٢٠١ | جبر المتمّمات المنتهية |
| Almost Certainly Event | ٢١٤ | حادث شبه أكيد |
| Almost Impossible Event | ٢١٠ | حادث شبه مستحيل |
| Almost Sure | ٢٣٧ | تقريباً أكيد |
| Alternative Hypothesis | ٥١٤ | الفرضية البديلة |
| Amplitude | ٥٩٢ | مطال |
| Analysis to Factors | ٤٦٩ | التحليل إلى عوامل |
| Anderson-Darling Test for Goodness of Fit | ٥٤٢ | اختبار أندرسون– دارلينغ لجودة التوفيق |
| Angular Frequency | ٥٩٢ | السرعة الزاوية للاهتزاز |
| Anscombe's Quartet | ١٣٢ | رباعية إنسكومب |
| Ascending Cumulative Frequency | ٣٢ | التكرار المتجمّع الصاعد |
| Ascending Cumulative Frequency Paragon (Less than Ogive) | ٤٢ | مضلّع التكرار المتجمّع الصاعد (أقل من قوسلقوطي) |
| Asymptotic Unbiased Estimator | ١٨٧ | مقدّر منصف تقاربياً |
| Asymptotically Normal Distributed | ٤١٩ | مقارب للتوزيع الطبيعي |
| Atom of the Partition | ٢٠٣ | ذرة تجزئة |
| Atomic Event | ١٩٥ | حادث ذري |
| Autoregressive Integrated Moving Average (ARIMA) | ٦٠٤ | الانحدار الذاتي والمتوسطات المتحركة المتكاملة |
| | **B** | |
| Bar Graph | ١٧ | العرض الشرائطي |
| Base Period | ٦١٥ | دورة الأساس |
| Bays's theorem | ٢٢١ | نظرية بييز |

| | | |
|---|---|---|
| Beppo – Levi's Theorem | ٦٦٢ | نظرية بيبو – ليفي |
| Bernoulli Distribution | ٣٠٣ | توزيع برنولي |
| Bernoulli Experiment | ٣٠٤ | تجربة برنولية |
| Bernoulli Random Variable | ٣٠٤ | متغيِّر عشوائي برنولي |
| Bernoulli's Law of Large Numbers | ٤١٦ | قانون برنولي للأعداد الكبيرة |
| Bertrand Paradox | ٢١١ | مُحيِّرة برتراند |
| Best Test | ٥١٩ | أفضل اختبار |
| Beta Function | ٦٤٩ | دالة بيتا |
| Binomial Distribution | ٢٨٢ | التوزيع الحدَّاني |
| Bochner Theorem | ٣٩٦ | نظرية بوخنر |
| Borel Field | ٦٥٣ | حقل بوريلي |
| Borel Set | ٢١٤ | مجموعة بوريلية |
| Bose – Einstein Statistic | ١٩٤ | إحصاءة بوزي – أينشتاين |
| Bowley's Coefficient of Skewness | ١٠٢ | مُعامل بولي للالتواء |
| Box plot | ١٠٧ | العرض (أو التمثيل) الصندوقي |
| | **C** | |
| Cardinality of a Set | ٦٤٢ | قدرة مجموعة |
| Cauchy – Schwartz inequality | ٣٦٤ | متباينة كوشي– شفارتز |
| Cauchy Distribution | ٣٣٦ | توزيع كوشي |
| Cauchy Sequence | ٤١٣ | متتالية كوشي |
| Cauchy–Schwartz Inequality | ٣٦٤ | متباينة كوشي– شفارتز |
| Causal Chain Relation | ١١٧ | علاقة السلسلة السببية |
| Causal Correlation | ١١٧ | الارتباط السببي |
| Celsius | ١٠ | وحدة لقياس الحرارة (الدرجة المئوية) |
| Centered Moment | ٣٧٢ | العزم المركزي |
| Central limit theorem's | ٤١١ | نظريات النهاية المركزية |
| Central Moment | ٩٦ | عزوم مركزي |
| Central Moment of Order r | ٩٦ | العزم المركزي من المرتبة r |
| Central Tendency Measure | ٥٦ | مقياس نزعة مركزيّة |
| Certain Event | ١٩٦ | الحادث الأكيد |
| Characteristic Function of PD | ٣٩٢ | الدالّة المميِّزة للتوزيعات الاحتمالية |
| Characteristic Function of RV | ٣٩٣ | الدالّة المميِّزة لمتغيِّر عشوائي |
| Chebyshev Inequality | ٣٨١ | متباينة تشبيشيف |
| Chebyshev Empirical Rule | ٨٩ | قاعدة تشبيشيف التجريبيّة |
| Chi Square Distribution | ٣٣٨ | توزيع كاي مربَّع |
| Chi-Squared Test for Goodness of Fit | ٣٣٧ | اختبار كاي مربَّع لجودة التوفيق |

| | | |
|---|---|---|
| Class (Group, Category or Interval) | ٢٤ | الفئة (في الجداول التكرارية) |
| Class Boundaries | ٢٧ | الحدود العمليّة للفئة |
| Class Limit | ٣٠ | حدود الفئة |
| Classes Table | ٢٤ | جدول تكراري ذو فئات |
| Classical Definition of Probability | ٢٠٧ | التعريف التقليدي للاحتمال |
| Classification of Samples | ٤٤١ | تصنيف العيّنات |
| Cluster Sample | ٤٤٦ | عيّنة عنقودية |
| Coefficient of Association | ١٤٠ | مُعامل الاقتران |
| Coefficient of Contingency | ١٤١ | مُعامل التوافق |
| Coefficient of Determination | ١٣٩ | مُعامل التحديد |
| Coefficient of Dispersion | ٩٣ | مُعامل التشتت |
| Coefficient of Variation | ٩٢ | مُعامل التّغيُّر |
| Coin | ١٩٠ | قطعة نقود معدنية |
| Combinations | ٦٤٥ | توافيق |
| Compact Set | ٢٣٩ | مجموعة متراصة |
| Compare Period | ٦٠٦ | دورة المقارنة |
| Complement of an Event | ١٩٦ | متمّم حادث |
| Complex Random Variable | ٢٣٨ | متغيِّر عشوائي عقدي |
| Composite hypothesis | ٥١٨ | فرضية مركّبة |
| Composite Index Number | ٦٢٣ | رقماً قياسياً مركّباً |
| Compound Binomial Distribution | ٢٨٢ | توزيع حداني مُركّب |
| Compound Distribution | ٢٨١ | توزيع مركّب |
| Comprehensive Survey | ٢ | المسح الشامل |
| Concave Function | ٦٤٧ | دالّة مقعّرة |
| Conditional Distribution | ٢٧٣ | توزيع شرطي |
| Conditional Expectation | ٣٩٨ | التوقع الشرطي |
| Conditional Probability | ٢١٨ | احتمال شرطي |
| Confidence Interval | ٤٩٢ | فترة ثقة |
| Confidence Level | ٤٩٣ | مستوى الثقة (أو معامل الثقة) |
| Confusing Variable | ١٧٨ | متغيِّر مشوّش |
| Consistency Property | ٤٨٨ | خاصيّة الاتساق |
| Consistent Estimator | ٤٨٨ | مقدِّر متّسق |
| Continuity theorem's | ٣٩٢ | نظريات الاستمرار |
| Continuous from Above | ٢١٥ | الاستمرار من الأعلى |
| Continuous from Below | ٢١٥ | الاستمرار من الأدنى |
| Continuous Probability Distribution | ٣٢٣ | توزيع احتمالي مستمرّ |

| | | |
|---|---|---|
| Continuous Random Variable | ٢٥٨ | متغيّر عشوائي مستمر |
| Continuous Random Vectors | ٢٧٣ | متّجه عشوائي مستمر |
| Continuous Set | ٤ | مجموعة مستمرّة (أو متصلة) |
| Continuous Uniform Distribution | ٣٢٣ | التوزيع المنتظم المستمر |
| Continuous Variable | ٦ | متغيّر مستمرّ |
| Convergence in Square Mean | ٤١٢ | التقارب بالمتوسط التربيعي |
| Converges Almost Sure | ٤١٢ | التقارب تقريباً أكيد (أو شبه أكيد) |
| Converges in Distribution | ٤١٢ | التقارب بالتوزيع |
| Converges in p-mean | ٤١٢ | التقارب بالمتوسط p |
| Converges in probability | ٤١٢ | التقارب بالاحتمال |
| Converges with probability equal to one | ٤١٢ | التقارب باحتمال يساوي الواحد |
| Convex Function | ٦٤٧ | دالّة محدَّبة |
| Convex Set | ٦٤٤ | مجموعة محدَّبة |
| Convolution theorem's | ٣٩٢ | نظريات التلاف |
| Correlation Coefficients | ١٢١ | مُعاملات الارتباط |
| Countable Set | ٦٤٢ | قابلة للعدّ |
| Covariance | ١٢٢ | التغاير |
| Covariance Matrix | ١٢٥ | مصفوفة التغاير |
| Critical Point | ٥١٦ | نقطة حرجة |
| Critical Region | ٥١٥ | منطقة حرجة |
| Critical Region of the Test | ٥١٥ | المنطقة الحرجة للاختبار |
| Critical Value | ٥١٦ | قيمة حرجة |
| Cumulative Distribution Function | ٢٤٣ | دالّة التوزيع التراكمية |
| Cumulative Frequency Paragon | ٤٢ | مضلّعات تكراري تراكمي |
| Current Period | ٦١٦ | الدورة الحالية (أو الدورة الجارية) |
| Curve | ٢١ | خط منحن |
| Curves Fitting | ١٤٣ | توفيق المنحنيات |
| | **D** | |
| Decision-Making Theory | ٥١٣ | نظرية اتخاذ القرار |
| Default Class width | ٢٨ | السعة الافتراضية (لفئة في جدول تكراري) |
| Default Number | ٢٨ | العدد الافتراضي (لفئات جدول تكراري) |
| Degenerate Distribution | ٣٠١ | التوزيع المضمحل |
| Degree | ٢٢ | درجة (تستخدم لقياس الزوايا) |
| Dependent Variable | ١١٥ | المتغيّر التابع |
| Descending Cumulative Frequency | ٣٢ | التكرار المتجمّع الهابط |

| | | |
|---|---|---|
| Descending Cumulative Frequency Polygon (Less than Ogive) | ٤٢ | مضلّع التكرار المتجمّع الهابط (أقل من قوس قوطي) |
| Descriptive Measures | ٥٥ | مقاييس وصفية |
| Dirac Measure | ٢١٥ | قياس ديراك |
| Discrete | ٤ | متقطّع (أو منفصل) |
| Discrete Probability Distribution | ٣٠١ | توزيع احتمالي متقطّع |
| Discrete Random Variable | ٢٥٢ | متغيّر عشوائي متقطّع |
| Discrete Random Vector | ٢٧٢ | متّجه عشوائي متقطّع |
| Discrete Uniform Distribution | ٣٠٧ | التوزيع المنتظم المتقطّع |
| Discrete Variable | ٦ | متغيّر متقطّع |
| Dispersion Measure | ٨٢ | مقياس تشتت |
| Distribution Law | ٢٤٠ | قانون توزيع |
| Dots Representation | ١٦ | تمثيل نقطي |
| Draw with Replacement | ٤٤٠ | السحب مع الإرجاع (أو مع الإعادة) |
| Draw without Replacement | ٤٤٠ | السحب بدون إرجاع (أو بدون إعادة) |
| | **E** | |
| Efficient Estimator | ٤٨٩ | مقدّر كفوء |
| Elementary Event | ١٩٥ | حادث ابتدائي |
| Empirical Distribution Function | ٤٥٣ | دالّة التوزيع العمليّة |
| Empirical Variance | ٨٦ | التباين العملي (أو تباين العيّنة) |
| Equal Almost Every Where | ٦٦٢ | يساوي تقريباً في كل مكان |
| Error | ١٤٥ | خطأ |
| Error of the First Type | ٥١٥ | الخطأ من النوع الأول |
| Error of the Second Type | ٥١٥ | الخطأ من النوع الثاني |
| Essentially Sequence | ٤١٣ | متتالية أساسية |
| Estimation Theory | ٤٧٧ | نظرية التقدير |
| Estimator more efficient from another | ٤٨٩ | المقدّر الأكفأ من مُقدّر آخر |
| Expected Frequencies | ٥٤٩ | التكرارات المتوقّعة |
| Explained Variation | ١٦٠ | التغيُّر المُفسَّر |
| Explanatory Variable | ١١٥ | متغيّر تفسيري |
| Exponential Distribution | ٣٢٤ | التوزيع الأسي |
| Exponential Regression | ١٦٨ | الانحدار الأسيّ |
| Extreme Value | ٧٧ | قيمة متطرفة |
| | **F** | |
| Factorial Moment | ٣٧٥ | العزوم العامليّ (أو الضربيّ) |
| Fahrenheit | ١٠ | وحدة الفهرنهايت لقياس الحرارة |

| | | |
|---|---|---|
| Failure Probability | ٣٠٤ | احتمال الفشل |
| Fermi-Dirac Statistic | ١٩٤ | إحصاءة فيرمي – ديراك |
| Finite Measure | ٦٥٧ | قياس منته |
| Finite Permutation | ٦٤٥ | تبديل منته |
| Finite Subadditive | ٦٥٧ | نصف جمعي منتهٍ |
| First Quartile (or Lower Quartile $Q_1$ ) | ٧٣ | الربيعي الأول (أو الربيعي الأدنى) |
| Fisher distribution | ٣٤٠ | توزيع فيشر |
| Fisher -Neumann Proposition | ٤٦٨ | مبرهنة فيشر – نويمان |
| Fisher's ideal index | ٦٣١ | المؤشّر المثالي ل فيشر |
| Fisher's Information Quantity | ٤٧١ | كمية معلومات فيشر |
| Fitted Value | ١٤٦ | قيمة موفّقة (تمّ توفيقها من منحنى التوفيق) |
| Fitting a Model of a Time Series | ٦٠٩ | توفيق نموذج متسلسلة زمنية |
| Five Number and the Box plot | ١٠٧ | الأعداد الخمسة والتمثيل الصندوقي للبيانات |
| Forecasting | ٦١١ | التنبؤ |
| Form Measures | ٩٦ | مقياس الشكل |
| Fourier Translation | ٣٩٣ | تحويل فورييه |
| Framework | ٤٣٩ | الإطار (إطار العمل) |
| Frequency | ١٣ | تكرار الفئات (عدد العناصر في فئة) |
| Frequency Curve | ٤٣ | منحنى تكراري |
| Frequency Distribution Table | ٣٠ | جدول توزيع تكراري |
| Frequency Histogram | ٣٦ | مدرّج تكراري |
| Frequency of Class | ٢٤ | تكرار الفئة |
| Frequency Polygon | ٤١ | مضلّع تكراري |
| Frequency Table | ١٣ | جدول تكراري |
| Fubini Theorem | ٦٦٢ | نظرية فوبيني |
| | G | |
| Gama Function | ٦٤٨ | دالة غاما |
| Geiger–Marsden experiment | ٣ | تجربة جايجر ومارسيدن |
| Geometric Definition of Probability | ٢١٠ | التعريف الهندسي للاحتمال |
| Geometric Distribution | ٣١٤ | التوزيع الهندسي |
| Geometric Mean | ٦٢ | المتوسط الهندسي |
| Gnedenko Problem | ٤٣٤ | مسألة غنيدينكو |
| Gnedenko Theorem | ٤٣٤ | نظرية غنيدينكو |
| Goodness of Distribution Fit Test | ٥٣٧ | اختبار جودة توفيق توزيع احتمالي |
| Graphing of Data | ١١ | عرض البيانات |
| Grouped Data | ٢٤ | بيانات مُجمّعة (بيانات مُجدولة) |

| | | |
|---|---|---|
| Growth Function | ٥٨٣ | دالّة النمو |
| | H | |
| Halberg's Cosine Method | ٥٩١ | طريقة التجيب ل هالبرغ |
| Harmonic Mean | ٦٤ | المتوسط التوافقي |
| Harmonic Model | ٥٧٠ | نموذج (أو نمط) توافقي |
| Head | ٤ | صورة (يقابلها الكتابة على قطع النقود المعدنية) |
| Highest Efficient Estimator | ٤٨٩ | المقدّر الأعلى كفاءة |
| Highest Fence (HF) | ٧٧ | الحد الأدنى للقيم المتطرفة بكبرها |
| High-Size Sample | ٤٤٧ | عيّنة كبيرة الحجم |
| Holt-Winters Smoothing | ٦٠٩ | التنعيم بطريقة هولت-وينتر |
| Homogeneity Test of Distribution | ٥٤٤ | اختبار التجانس لتوزيع احتمالي |
| Hypergeometric Distribution | ٣١٨ | التوزيع فوق الهندسي (أو الفوهندسي) |
| | I | |
| Impossible Event | ١٩٦ | الحادث المستحيل |
| Imprint of Time Series | ٥٦٥ | بصمة متسلسلة زمنية |
| Independence of Events | ٢٢٣ | استقلال حوادث |
| Independence Test of two Phenomenon | ٥٤٨ | اختبار استقلال ظاهرتين |
| Independent | ٢٢٣ | مستقل |
| Independent Events | ٢٢٣ | حوادث مستقلّة |
| Independent Populations | ٤٦١ | مجتمعات مستقلّة |
| Independence of Random Variables | ٣٥١ | استقلال المتغيّرات العشوائية |
| Independent Random Variables | ٣٥١ | متغيّرات عشوائية مستقلّة |
| Independent Variables | ٣٥١ | متغيّرات مستقلّة |
| Index Number | ٦١٨ | رقم قياسي |
| Indicator Function | ٦٤٦ | الدالّة المميّزة |
| Indicator Function of a Set | ٦٤٦ | الدالّة المميّزة لمجموعة |
| Indicator Function of an Event | ٢٣٩ | الدالّة المميّزة لحادث |
| Initial Element | ٤٤٥ | عنصر البدء |
| Integrable | ٢٥٩ | كمول (أو قابل للمكاملة) |
| Integrablity of a Real Function | ٦٦١ | قابلية المكاملة لدالّة حقيقية |
| Integral of a Measurable Function | ٦٦١ | تكامل دالّة قيوسة |
| Intentional Samples | ٤٤١ | عيّنات عمديّة (أو قصديّة) |
| Interquartile Range (IQR) | ٩١ | المدى الربيعي |
| Interval Scale | ٩ | مقياس الفترة |
| Inversion theorems | ٣٩٢ | نظريات العكس |
| Irregular Component | ٥٧٠ | مُركّبة عشوائية (أو غير منتظمة) |

| | | |
|---|---|---|
| Irregular Cyclical | ٥٧٠ | دورية غير منتظمة (عشوائية) |
| | J | |
| Jensen inequality | ٣٦٣ | متباينة جينسين |
| Joint Frequency Table | ١٢٣ | الجدول التكراري المشترك |
| | K | |
| Kelvin | ١٠ | وحدة الكلفين لقياس الحرارة المنخفضة جداً |
| Kolmogorov-Smirnov Test | ٥٣٩ | اختبار كلموغوراف – سميرنوف |
| Kurtosis Measure | ١٠٥ | مقياس تفلطح (أو تفرطح) |
| | L | |
| Lag | ٥٦٤ | فجوة زمنية |
| Laspeyres Index | ٦٢٥ | مؤشّر لسبيريس |
| Laws of Large Numbers | ٤١٤ | قوانين الأعداد الكبيرة |
| Least Square Method | ١٤٤ | طريقة المربعات الصغرى |
| Lebesgue Measure | ٦٥٧ | قياس لوبيغ |
| Length of Class | ٢٤ | سعة الفئة (في جدول تكراري) |
| Length of Period | ٥٩٢ | طول الدور |
| Leptokurtic | ٤٨ | مُدبَّب |
| Leptokurtic Distribution | ٤٨ | توزيع مدبَّب |
| Level | ٦٠٩ | المستوى (أو التمهيد) |
| Lévy Theorem | ٣٩٧ | نظرية ليفي |
| Likelihood Function | ٤٨١ | دالّة الأرجحية |
| Lindenberg – Feller Problem | ٤٣١ | مسألة ليندنبرغ – فيلر |
| Lindenberg – Feller Theorem | ٤٣١ | نظرية ليندنبرغ – فيلر |
| Lindenberg – Levy Problem | ٤٢١ | مسألة ليندنبرغ – ليفي |
| Lindenberg–Levy Theorem | ٤٢٢ | نظرية ليندنبرغ– ليفي |
| Local Central Limit Theorem | ٤٣٤ | نظرية النهاية المركزية الموضعية |
| Logarithmic Regression | ١٦٧ | الانحدار اللوغاريتمي |
| Logistic Function | ٥٨٤ | الدالّة المنطقية |
| Lower Confidence Limit | ٤٩٣ | حد الثقة الأدنى |
| Lower Continuity of a measure | ٦٥٨ | الاستمرار من الأدنى لقياس |
| Lowest Fence (LF) | ٧٧ | الحد الأعلى للقيم المتطرفة بصغرها |
| Lyapunov Inequality | ٣٧١ | متباينة ليابانوف |
| Lyapunov Problem | ٤٢٨ | مسألة ليابانوف |
| Lyapunov Theorem | ٤٢٨ | نظرية ليابانوف |
| | M | |
| Marginal Density Function | ٢٧٨ | دالّة الكثافة الهامشية |

| | | |
|---|---|---|
| Marginal Distribution | ٢٧٨ | توزيع هامشي |
| Marginal Distribution Function | ٢٧٩ | دالّة التوزيع الهامشية |
| Markov Condition | ٤٨٨ | شرط ماركوف |
| Markov inequality | ٣٧١ | متباينة ماركوف |
| Martingale | ٣٩٨ | الحَكَمة |
| Martingale Theorem | ٣٩٨ | نظرية الحَكَمة |
| Mathematical Expectation | ٣٥٩ | التوقع الرياضياتي |
| Mathematical Statistics | ١٨٨ | الإحصاء الرياضياتي |
| Maximum Likelihood Estimator | ٤٨١ | مُقدِّر الأرجحية العظمى |
| Maximum Likelihood Method | ٤٧٨ | طريقة الأرجحية العظمى |
| Maxwell–Boltzmann Statistic | ١٩٤ | إحصاءة ماكسويل– بولتسمان |
| Mean (or Arithmetic Mean) | ٥٦ | المتوسط (أو المتوسط الحسابي) |
| Mean Absolute Deviation (MAD) | ٦١١ | متوسط القيم المطلقة للانحرافات |
| Mean Absolute Percentage Error (MAPE) | ٦١١ | متوسط القيم المطلقة للنسب المنوية للأخطاء |
| Mean of Sample | ٤٥٢ | متوسط العيِّنة |
| Mean Square Deviation (MSD) | ٦١١ | متوسط مربَّع الانحرافات |
| Mean Squared Error | ١٦٠ | متوسط مربَّع الخطأ |
| Mean Deviation (or Average Deviations) | ٨٣ | الانحراف المتوسط (أو متوسط الانحرافات) |
| Measurable Function | ٦٥٥ | دالّة قيوسة (أو قابلة للقياس) |
| Measurable Map | ٦٥٥ | تطبيق قيوس (أو قابل للقياس) |
| Measurable Space | ٢١٤ | فضاء مقيس |
| Measure | ٦٥٦ | قياس |
| Measure Space | ٦٥٦ | فضاء قياس |
| Measure Theory | ١٥٩ | نظرية القياس |
| Measuring Fitting Accuracy | ٦١١ | قياس جودة التوفيق |
| Median | ٦٦ | وسيط (أو الوسط) |
| Mesokurtic Distribution | ٤٨ | توزيع معتدل |
| Methods of Sampling | ٤٤٠ | طرائق سحب العيِّنة |
| Metric Space | ٦٥٣ | فضاء متري |
| Midpoint | ٣١ | مركز الفئة (في جدول تكراري) |
| Mixed Random Vector | ٢٧٤ | متَّجه عشوائي مختلط |
| Mode | ٧٠ | المنوال |
| Moivre – Laplace Problem | ٤١٧ | مسألة موافير– لابلاس |
| Moment Coefficient of Skewness | ١٠٠ | مُعامل الالتواء العزومي |
| Moment Generating Function | ٣٨٩ | الدالّة المولِّدة للعزوم |
| Moments Method | ٤٧٨ | طريقة العزوم |

| | | |
|---|---|---|
| Moments of random variable | ٣٥٨ | العزوم لمتغيِّر عشوائي |
| Monetary Units | ٦١٩ | وحدة نقدية |
| Monotone class | ٦٥٤ | صف مطَّرد |
| Monotony | ٦٥٧ | الاطّراد |
| Moving Average | ٦٠٢ | متوسط متحرك |
| Multimodal Symmetric Distribution | ٤٩ | توزيع متناظر متعدّد المنوال |
| Multiple Correlation | ١١٧ | الارتباط المتعدّد |
| Multiplication Law of Probability | ٢٢٠ | قانون الضرب في الاحتمالات |
| Multiplication Rule | ٦٤٦ | قاعدة الضرب |
| Multiplicative Model | ٥٧٢ | نموذج الضرب |
| Mutually Exclusive Events | ١٩٧ | حوادث متنافية |
| | **N** | |
| Negative Binomial Distribution | ٣١١ | التوزيع الحدَّاني السالب |
| Negative Skewed Distribution | ٥٠ | توزيع سالب الالتواء |
| Nets Distribution | ٤٣٤ | توزيعات شبكية |
| Neumann Deference Test | ٥٨٧ | اختبار الفرق لـ نويمان |
| Neyman Pearson Theorem | ٥١٩ | نظرية نيمان – بيرسون |
| Nominal Scale | ٧ | مقياس اسمي |
| Non Linear Correlation | ١١٨ | الارتباط غير الخطي |
| Non Parametric Test | 513 | اختبار غير مَعْلمي |
| Non Symmetric Distributions | ٤٩ | توزيع تكراري غير متناظر |
| Normal (or Gaussian) Distribution | ٣٢٧ | التوزيع الطبيعي (أو التوزيع الغاوسي) |
| Normalized random variable | ٣٧٥ | متغيِّر عشوائي مستنظم |
| Null Hypothesis | ٥١٤ | الفرضية الابتدائية (أو الفرضية الصفرية) |
| Number of Classes | ٣٠ | عدد الفئات |
| | **O** | |
| Observation | ٣ | ملحوظة (أو مشاهدة) |
| Observation Values | ٤٤٩ | قيم ملحوظة (أو مشاهدة) |
| Ogive | ٤٣ | القوس القوطي |
| One Point Distribution | ٣٠١ | التوزيع وحيد النقطة |
| One -Tailed Test | ٥١٨ | اختبار أحادي الطرف (أو أحادي الذيل) |
| Optimal Test | ٥٢٠ | الاختبار الأمثل |
| Optimum Rejection Area | ٥١٩ | منطقة الرفض المثلى |
| Order of the Partial Correlation | ١٧٨ | رتبة الارتباط الجزئي |
| Ordinal Scale | ٨ | مقياس ترتيبي |
| Outcome | ١٨٩ | نتيجة (نتيجة تجربة) |

| | | |
|---|---|---|
| Outlier Value | ٧٧ | قيمة منعزلة |
| | **P** | |
| Paasche index | ٦٢٥ | مؤشّر باشّي |
| Pairwise Independent | ٢٢٤ | مستقلّة مثنى مثنى |
| Pairwise Mutually Exclusive Events | ١٩٧ | حوادث متنافية مثنى مثنى |
| Parameter | ٥٦ | مَعْلَمة (أو وسيط) |
| Parametric Test | ٥١٣ | اختبار مَعْلمي |
| Partial Correlation | ١١٧ | الارتباط الجزئي |
| Partial Correlation Coefficient | ١٧٨ | مُعامل الارتباط الجزئي |
| Partition of a Set | ٢٠٣ | تجزئة مجموعة |
| Pauli Principle | ١٩٤ | مبدأ باولي (مبدأ الاستبعاد) |
| Percent Frequency | ٣٢ | التكرار المئوي |
| Percentile (Percentiles) | ٧٧ | مئين (المئينات) |
| Percentile Frequency Histogram | ٣٨ | مدرّج تكراري مئوي |
| Period (or Cyclical) Component | ٥٧٠ | المُركّبة الدورية |
| Permutation | ٦٤٥ | تباديل |
| Person's Coefficient of Correlation | ١٢٧ | مُعامل بيرسون للارتباط |
| Person's Coefficient of Kurtosis | ١٠٦ | مُعامل بيرسون للتفلطح |
| Person's Coefficient of Skewness | ١٠٠ | مُعامل بيرسون للالتواء |
| Perturbation Parameter | ٥١٨ | مَعْلَمة مشوّشة |
| Perturbation Variable | ١٧٨ | متغيّر مشوّش |
| Pie Chart | ٢١ | تمثيل للبيانات بقرص دائري |
| Platykurtic Distribution | ٤٨ | توزيع منبسط |
| Playing Carts | ٢٠٨ | بطاقات لعب |
| Point Estimation | ٤٨٥ | التقدير بنقطة (أو التقدير النقطي) |
| Point Process | ٢٣٨ | طوري نقطي |
| Poisson Distribution | ٢٨٢ | توزيع بواسون |
| Poisson's Law of Large Numbers | ٤١٦ | قانون بواسون للأعداد الكبيرة |
| Polonium | ٥٣٨ | البولونيوم (مادة مشعّة) |
| Polygon | ١٩ | مضلّع (تمثيل بقطع مستقيمة متتالية) |
| Pool algebra | ١٩٨ | جبر بول |
| Pooled Estimator of Variance | ٤٦٣ | المقدّر المشترك للتباين |
| Population | ١ | مجتمع إحصائي |
| Population Parameter | ٤٧٧ | مَعْلَمة مجتمع إحصائي |
| Position Measure | ٥٥ | مقياس موضع |
| Positive Definite | ٣٩٤ | موجبة التعيين |

| | | |
|---|---|---|
| Positive Skewed Distribution | ٤٩ | توزيع موجب الالتواء |
| Power of the Test | ٥١٥ | قوة الاختبار (أو حساسية الاختبار) |
| Power Regression | ١٦٩ | انحدار القوى |
| Predicted Value | ١٤٦ | قيمة مقدَّرة |
| Predictor Variable | ١٤٦ | المتغيِّر المُتنبأ به |
| Previous Cycle | ٦١١ | دورة سابقة |
| Prior Probability | ٢٢١ | الاحتمال الأسبق |
| Probability Generating Function | ٣٨٣ | الدالّة المولِّدة الاحتمالية |
| Probability Mass | ٢٥٣ | كتلة احتمالية |
| Probability Mass Function | ٢٥٣ | دالّة الكتلة الاحتمالية |
| Probability Measure | ١٨٩ | قياس احتمالي |
| Probability Space | ١٨٧ | فضاء احتمالي |
| Probability Theory | ١٨٨ | نظرية الاحتمالات |
| Projection Function | ٢٦٦ | دالّة الإسقاط |
| Pulling with Replacement | ١٩٢ | السحب مع الإرجاع |
| Pulling without Replacement | ١٩٣ | السحب بدون إرجاع |
| | **Q** | |
| Qualitative Data | ٦ | بيانات نوعيّة |
| Qualitative Variables | ٧ | متغيِّرات نوعيّة |
| Quantitative Data | ٩ | بيانات كميّة |
| Quantitative Variable | ٧ | متغيِّر كميّ |
| Quartile (Quartiles) | ٧٣ | ربيعي (ربيعيات) |
| Queuing theory | ٣٢٠ | نظرية الطوابير |
| Quota Sample | ٤٧٧ | عيِّنة حصص (مبنية على المحاصصة) |
| | **R** | |
| Radon Nikodym Density Function | ٢٥٨ | دالّة الكثافة ل رادون- نيكوديم |
| Radon–Nikodym Theorem | ٦٣٣ | نظرية رادون – نيكوديم |
| Random Closed Set | ٢٣٩ | مجموعة مغلقة عشوائية |
| Random Element | ٢٣٨ | عنصر عشوائي |
| Random Error Component | ١٥٥ | مركبة الخطأ العشوائي |
| Random Experiment | ١٨٨ | تجارب عشوائية |
| Random Kern | ٢٣٩ | نواة عشوائية |
| Random Process | ٥٦٤ | طوري عشوائي |
| Random Sample | ٤٤١ | عيِّنة عشوائية |
| Random Variable | ٢٣٣ | متغيِّر عشوائي |
| Random Vector | ٢٣٨ | متَّجه عشوائي |

| | | |
|---|---|---|
| Random Walk | ٢٣١ | الخطى العشوائية |
| Range | ٢٧ | المدى |
| Rank | ٧٤ | رتبة |
| Rao-Cramer Inequality | ٤٩٠ | متباينة راو- كرامر |
| Ratio to Moving-Average Method | ٦٠٥ | طريقة النسبة إلى متوسط متحرك |
| Ration Scale | ٩ | مقياس النسبة |
| Raw Data | ١٢ | بيانات خام |
| Real Random Process | ٥٦٤ | طوري عشوائي حقيقي |
| Reciprocal Correlation | ١١٧ | الارتباط التبادلي |
| Reciprocal Relation | ١١٧ | علاقة تبادلية |
| Regression Coefficients | ١٤٦ | مُعاملات (أو وسطاء) الانحدار |
| Regression Equation | ٤٠٤ | معادلة انحدار |
| Regression Line | ٤٠٤ | خط انحدار |
| Regular Experiment | ١٨٧ | تجارب نظامية |
| Regular Random variable | ٤٧١ | متغيِّر عشوائي نظامي |
| Rejection Area | ٥١٥ | منطقة الرفض |
| Relative Efficiency | ٤٨٩ | الكفاءة النسبية |
| Relative Frequency | ٣١ | التكرار النسبي |
| Relative Frequency Histogram | ٣٧ | مدرَّج التكراري النسبي |
| Relative Standard Deviation | ٩٣ | الانحراف المعياري النسبي |
| Relative Value | ٦١٨ | قيمة نسبية |
| Reliability Theory | ٣٢٠ | نظرية الوثوقية |
| Representation by Picture | ٢٣ | التمثيل بالصور (الطريقة التصويرية) |
| Residual | ١٤٥ | الباقي |
| Resistance | ٤ | مقاومة أوميَّة |
| Response Variable | ١١٥ | متغيِّر الاستجابة |
| Richest Algebra (or Strongest Algebra) | ٢٠٢ | أغنى جبر (أو أقوي جبر) |
| Roughness Coefficient | ٦٠١ | مُعامل الخشونة |
| $r$-Percentile (or $r^{th}$ Percentile) | ٧٧ | المئين الرائي |
| | S | |
| Sample | ٢ | عيِّنة |
| Sample Function | ٤٥١ | دالَّة عيِّنة |
| Sampling Distributions | ٤٥٦ | توزيعات المعاينة |
| Sampling Theory | ٣ | نظرية المعاينة |
| Sampling Unit | ٤٣٩ | وحْدَة المعاينة |
| Satterthwaite's Approximation | ٤٦٣ | تقريب ستيرثويت |

| | | |
|---|---|---|
| Scatter Measure | ٨٣ | مقياس التبعثر(تسمية أخرى لمقياس التشتت) |
| Scatter Plot | ١١٩ | لوحة الانتشار |
| Seasonal Changes | ٦٠٤ | التغيُّرات الموسميّة |
| Seasonal Component | ٥٦٩ | المُركّبة الموسميّة |
| Seasonal Cycle | ٥٧٣ | الدورة الموسميّة |
| Seasonal Parameter | ٦١٠ | مَعْلَمَة الموسمي |
| Second Quartile | ٧٣ | الربيعي الثاني |
| Selected Sample | ٤٤٧ | عيّنات منتقاة |
| Separable | ٢٣٨ | فَصول (أو قابل للفصل) |
| Set of Outcomes | ١٨٩ | مجموعة نتائج |
| Significance Level | ٥١٥ | مستوى أهمية |
| Significance Level of the Test | ٥١٥ | مستوى أهمية الاختبار |
| Simple Aggregate Index Number | ٦٢٣ | رقم قياسي كلي بسيط |
| Simple Average | ٦١٩ | المعدّل البسيط |
| Simple Correlation | ١١٧ | الارتباط البسيط |
| Simple Event | ١٩٦ | حادثٌ بسيط |
| Simple Fractional Regression | ١٦٧ | الانحدار الكسري البسيط |
| Simple Function | ٦٤٧ | دالّة بسيطة |
| Simple hypothesis | ٥١٨ | فرضية بسيطة |
| Simple Index Number | ٦١٨ | رقم قياسي بسيط |
| Simple Linear Correlation | ١١٥ | الارتباط الخطي البسيط |
| Simple Random Sample | ٤٤٤ | العيّنة العشوائية البسيطة |
| Size of Population | ٢ | حجم المجتمع |
| Size of Sample | ٣ | حجم العيّنة |
| Skewed Frequency Distribution | ٤٩ | توزيع تكراري ملتوي |
| Skewed to the Left Distribution | ٥٠ | توزيع ملتوي نحو اليسار |
| Skewed to the Right Distribution | ٤٩ | توزيع ملتوي نحو اليمين |
| Skewness Measure | ٩٩ | مقياس التواء |
| Smoothing Model | ٥٧٢ | نموذج تنعيم |
| Smoothing of Time Series | ٦٠١ | تنعيم متسلسلات زمنية |
| Space of Elementary Events | ١٩٥ | فضاء الحوادث الابتدائية |
| Spearman Rho Correlation Coefficient | ١٣١ | مُعامل سبيرمان رو للارتباط |
| Spearman's Rank Correlation Coefficient | ١٣٠ | مُعامل سبيرمان لارتباط الرتب |
| Spline | ٢١ | التمهيد الشرائحي |
| Spread Measures | ٨٣ | مقياس الانتشار(تسمية أخرى لمقياس التشتت) |
| Spurious Correlation | ١١٧ | الارتباط الوهمي |

| | | |
|---|---|---|
| Spurious Relation | ١١٧ | علاقة وهمية |
| Standard Deviation | ٨٧ | الانحراف المعياري |
| Standard Error | ١٥٨ | الخطأ المعياري |
| Standard Error of the Estimate | ١٥٨ | الخطأ المعياري للتقدير |
| Standard Score | ٩٣ | درجة معيارية |
| Standardization | ٩٥ | استعيار أو استقياس (أي جعله معيارياً أو قياسياً) |
| Statistic | ٤٥١ | إحصاءة |
| Statistical Estimator | ٤٧٨ | مقدّر إحصائي |
| Statistical hypothesis | ٥١٣ | فرضية إحصائية |
| Statistical Hypothesis test | ٥١٣ | اختبار فرضية إحصائية |
| Statistical Independent | ٢٢٤ | مستقلّة إحصائياً |
| Statistical Inference | ٤٧٧ | الاستدلال الإحصائي |
| Staurt-Cox Test | ٦٨٨ | اختبار ستراوت-كوكس |
| Statistics | ١ | علم الإحصاء |
| Steiner Formula | ٣٧٤ | صيغة شتاينر |
| Stem and Leafs Table | ١٤ | جدول الساق والأوراق |
| Step Function | ٦٤٧ | دالّة درجيّة |
| Stirling Expansion | ٦٤٥ | نشر سترلنغ |
| Stochastic | ١٨٨ | علم العشوائيات |
| Stochastic Geometry | ٢١٣ | الهندسة العشوائية |
| Stochastic Independent | ٢٢٤ | مستقلّة عشوائياً |
| Straight Regression | ١٤٥ | مستقيم الانحدار |
| Stratified Sample | ٤٤٥ | عيّنة طبقية |
| Strong Law of Large Numbers | ٤١٤ | القانون القوي للأعداد الكبيرة |
| Student Distribution (or t-Distribution) | ٣٣٥ | توزيع ستودنت (أو توزيع t) |
| Subadditive | ٦٥٧ | نصف جمعي |
| Success Probability | ٣٠٤ | احتمال النجاح |
| Successive Differences Method | ٥٧٨ | طريقة الفروق المتتالية |
| Sufficiency Statistic | ٤٦٧ | الإحصاءة الكافية |
| Superior Probability | ٢٢١ | الاحتمال اللاحق |
| Symmetric Frequency Distribution | ٤٨ | توزيع تكراري متناظر |
| Systematic Component | ١٥٥ | المُركّبة النظامية (أو الرتيبة) |
| Systematic Sample | ٤٤٥ | العيّنة الرتيبة (أو المنتظمة) |
| Systematic Vibrations | ٥٦٩ | اهتزازات رتيبة |
| | **T** | |
| Tables Method | ١٣ | طريقة الجداول |

| | | |
|---|---|---|
| Tail | ١٩٠ | شعار (في تجربة قذف قطعة نقود معدنية) |
| Tally | ٣١ | التعداد (تعداد عناصر فئة في جدول تكراري) |
| Taylor's Expansion | ٣٩٠ | نشر تايلور |
| The Generated Partition by RV | ٣٩٩ | التجزئة المولّدة من متغيّر عشوائي |
| The $k^{th}$ Factorial Moment | ٣٧٦ | العزم العاملي من المرتبة k |
| The Risk of Type II | ٥١٥ | المجازفة من النوع الثاني |
| Theoretical Mathematical Model | ٤٥٠ | نموذج رياضياتي نظري |
| Third Quartile (or Upper Quartile $Q_3$) | ٧٣ | الربيعي الثالث (أو الربيعي الأعلى) |
| Time Series | ٥٦٣ | متسلسلة زمنية |
| Total Probability Formula | ٢٢٠ | صيغة الاحتمال التام |
| Total Variation | ١٦٠ | التغيُّر الكلي |
| Trend | ٥٦٩ | الاتجاه العام |
| Trend Component | ٥٦٩ | مُركّبة الاتجاه العام |
| Trend Parameter | ٦١٠ | مَعْلَمَة الاتجاه العام |
| Trimmed mean (TrMean) | ٧٢ | المتوسط الحسابي المشذّب |
| Trivial Algebra | ٢٠٢ | الجبر المبتذل |
| Trivial Partition | ٢٠٣ | التجزئة المبتذلة |
| $\tau$-Score | ٩٥ | الدرجة المعيارية -$\tau$ (الدرجة المعيارية التائية) |
| Tow Point Distribution | ٣٠٣ | التوزيع ثنائي النقطة |
| Two-Tailed Test | ٥١٨ | اختبار ثنائي الطرف (أو ثنائي الذيل) |
| | **U** | |
| Unbiased Estimator | ٤٨٦ | المقدّر المُنصِف (أو غير المنحاز) |
| Unbiasedness | ٤٨٦ | منصف أو غير متحيز |
| Unbiasedness Quantity | ٤٨٦ | مقدار التحيُّز |
| Uncorrelated | ٣٦٨ | غير مرتبطين |
| Uncountable Set | ٦٤٢ | غير قابلة للعدّ |
| Unexplained Variation | ١٦٠ | التغيُّر غير المُفسَّر |
| Unidirectional Relation | ١١٦ | علاقة وحيد الاتجاه |
| Unimodal Symmetric Distribution | ٤٨ | توزيع متناظر أحادي المنوال |
| Uniqueness theorem's | ٣٩٢ | نظريات الوحدانية |
| Upper Confidence Limit | ٤٩٣ | حد الثقة الأعلى |
| Upper Continuity | ٦٥٨ | الاستمرار من الأعلى |
| | **V** | |
| Variability Measure | ٨٣ | مقياس الاختلاف (تسمية أخرى لمقياس التشتت) |
| Variable | ٥ | متغيِّر |
| Variance | ٨٥ | التباين |

| | | |
|---|---|---|
| **Variance of Sample** | ٤٥٢ | تباين العيّنة |
| | **W** | |
| **Wallis –Moore Test** | ٥٨٨ | اختبار الفروق الأولى لووليس- مور |
| **Weak Law of Large Numbers** | ٤١٤ | القانون الضعيف للأعداد الكبيرة |
| **Weibull Distribution** | ٥٤٢ | توزيع وايبول |
| **Weight for the level** | ٦٠٩ | الوزن من أجل المستوى (أو من أجل التمهيد) |
| **Weight for the seasonal component** | ٦٠٩ | الوزن من أجل المركّبة الموسميّة |
| **Weight for the Trend** | ٦٠٩ | الوزن من أجل الاتجاه العام |
| **Weighted Geometric Mean** | ٦٣ | المتوسط الهندسي الموزون |
| **Weighted Index Numbers** | ٦٢٥ | رقم قياسي موزون |
| **Weighted Mean** | ٦١ | المتوسط الموزون |
| **Welch–Satterthwaite Equation** | ٤٦٣ | معادلة ويلش ستيرثويت |
| **Whiskers** | ١٠٨ | الشُعيْرتان (شُعيرتا التمثيل الصندوقي للبيانات) |
| **White Noise** | ٥٧٠ | الضجيج الأبيض |
| **White Noise Process** | ٥٧٠ | طوري ضجيج أبيض |
| **Winters' Method** | ٦٠٠ | طريقة وينتر |
| **Winter's Method for Generate Forecasts** | ٦٠٨ | طريقة وينتر لتوليد التنبؤات |
| | **Z** | |
| ***z*-Score** | ٩٤ | الدرجة المعيارية -ح |
| | **σ** | |
| **σ-Finite Measure** | ٦٥٧ | قياس σ- منتهٍ |
| **σ–algebra** | ١٩٩ | جبر من نوع σ (σ- جبر) |
| **σ–Finite** | ٦٥٧ | نصف جمعي σ- منتهٍ (أو نصف جمعي عدود) |